Essays on Automation, Inequality, and Intergenerational Mobility

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec
(corresponds to Doctor of Philosophy, PhD)

presented by

Carlo Zanella
from Hölstein, BL

approved in February 2020 at the request of
Prof. Dr. David Hémous
Prof. Dr. Florian Scheuer

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 12.02.2020

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

I would like to extend my deepest gratitude to my thesis supervisor and co-author David Hémous for his invaluable input, advice, and suggestions as well as guiding me through the program over the last two and a half years. I would also like to thank my co-author and unofficial mentor Winfried Koeniger. Winfried has supported me in every way possible ever since I began my Master's degree in St. Gallen, and it was him who encouraged me to do a PhD in the first place. *Thank you* to both of you.

I would also like to thank the Department of Economics at the University of Zurich for financial support during my graduate studies. I would like to thank Mirjam Britschgi, Karin Wyss, and Reto Frey for their kind support in overcoming administrative challenges during my studies.

My deepest gratitude to Florian Scheuer for being a second reader of this thesis and, more significantly, for the privilege to be his research and teaching assistant. I have learnt a great deal about macroeconomics (and Fréchet derivatives!) from working with him. I would also like thank all the graduate students whom I was teaching and who put up with my handwriting in the (at times probably chaotic) macro tutorials. I hope most of you could recover from this traumatic experience.

Further, I would like to express my gratitude to my fellow graduate students and friends for inspiring conversations, companionship, and friendship. I would especially like to thank Vladimir Sulaja for his friendship and the honest conversations about economic research. I would also like to thank Dóra Simon and Julia Grünseis who have both been an incredible enrichment to my life. Moreover, I would like to thank Ulrich Bergmann, Stefanie Bosshard, Caroline Chuard, Emanuele Dicarlo, Elisa Macchi, Michael Stiefel, Juliette Thibaud, Kremena Valkanova, and Tong

Zhang, for accompanying me during my journey.

I would also like to thank my friends outside of the economics profession. In particular, I would like to express my sincerest gratitude to Simon Marti for inspiring conversations as well as deep explorations into the nature of everything and nothing.

Finally, I want to thank my sister Patricia Zanella, my father Peter Zanella, my grandmother Margrith Tanner, as well as Christina Peikert for their unconditional, unfailing support throughout my whole life.

Contents

List of Figures

List of Tables

Overview

The last four decades have seen a rise in economic inequality throughout the world. This has sparked interest into distributional issues both in the public as well as in the economic profession. Several causes have been proposed for this increase in economic inequality.

One cause which frequently discussed in public debates are the recent advances in automation technology: Industrial robots as well as computerization have taken over tasks that have in the past been performed by humans. This technological change has greatly affected the workplace, and there are growing concerns that automation technologies displace workers, lead to higher unemployment, decrease wages, and eventually lead to higher inequality.

A second central issue underlying the question of inequality is about *equal opportunity*. Whereas inequality in incomes across individuals may be tolerated to some extent, many consider inequality in opportunities to be unfair. We have recently witnessed a significant concentration of wealth in the hands of a few. In the form of inheritances and investment into the human capital of children by their parents such inequality is transmitted from one generation to the next. This may lead children of rich parents to have better opportunities than children of poor families. It is clear that questions of social mobility and opportunity lie at the heart of the conerns about high economic inequality. This disseration studies these two topics of inequality in three chapters, which are briefly outlined below.

The first Chapter, entitled *Automating Labor: Evidence from Firm-level Patent Data*, studies to what extent increases in low-skilled wages lead firms to invent new automation technologies. Answering this question is essential to understand the long-term effects of policy inverventions such as higher minimum wages. Previous research has mostly focused either on automation *adoption* rather than *innovation* or on the reverse effects of automation innovation on wages. We are the first to study the effect of wages on automation innovation. In this chapter, we make two contributions. First, we develop a novel patent-based measure of automation innovation. To do that we classify patents as automation by using a combination of full-text keyword search and technological classification codes which are assigned to patents. We show that our measure is highly correlated with decreases in routine tasks in U.S. industries. Our second contribution is to use this patent classification to identify the causal effect of low-skilled wages on automation innovation. We identify the causal effect of wages on automation by building on the method by Aghion et al. (2016), which exploits firm-level variation in firms' exposure to countries as an identification strategy. We find substantial effects of wages on automation innovation: higher low-skill wages lead to more automation innovations with an elasticity which we estimate between 1 and 2.2 depending on our specification. Finally, we conduct an event study on the German Harzt reform 2002–2004 which was designed to increase labor flexibility and thereby decrease costs of labor. We find that foreign firms exposed to Germany sharply decreased innovation in automation technologies after the reform.

In the second Chapter, entitled *Automation and the Labor Share: Evidence from Patents*, we contribute to the understanding of the global decline in the labor share and relate it to automation. The labor share — that is the share of aggregate income paid out to labor in the form of wages, salaries, and benefits — has been decreasing steadily over the majority of countries in the last three decades.(Karabarbounis and Neiman, 2014) Given that most people make a living by selling their labor, this decline has vast implications for inequality. In our study, we employ the patent classification developed in Chapter 1 and document a robust negative relationship between automation and the labor share. We find that the negative association is mostly driven by the middle-skilled labor share. Moreover, automation innovation is negatively associated with the share of hours worked of low-skilled labor, not significantly related to the share of hours worked of middle-skilled labor, and positively related to the share of hours worked of high-skilled labor.

Chapter 1 and 2 both address the progress in automation technologies raised in the beginning of this section. While Chapter 2 shows how the development of automation is related to important global processes affecting inequality, Chapter 1 sheds light on the micro-level mechanisms that determine the innovation of automation technologies. I believe that Chapter 1 makes clear that automation is not an exogenous process and its relationship with inequality is two-sided: The attempt to mitigate inequality, for example, by increasing minimum wages will lead to more automation innovation, which may displace labor and eventually lead to more inequality.

Finally, in the third Chapter, entitled *Opportunity and Inequality Across Generations*, we address the second topic of this thesis: How important is parental background for the transmission of inequality? We study this question in an OLG model calibrated to the U.S. economy. Our quantitative results suggest that the transmission of inequality is determined mostly by nature (ability) rather than nurture (bequests and schooling). We then solve for the socially optimal allocation in this economy. In the social optimum, the importance of nurture increases relative to nature compared with the steady state of the calibrated economy. The planner provides more intergenerational insurance against ability risk at the cost of less mobility. However, this decrease in mobility is modest compared to the increase in insurance the planner is able to provide. The socially optimal allocation is highly non-linear and history-dependent in a manner that cannot easily be captured by current observables such as income, bequests, or schooling expenditures. We therefore assess the extent to which the welfare gains of the social optimum can be achieved using simple, history-independent tax schedules. We show that a linear tax system can capture almost half of the welfare gains of the socially optimal allocation.

Chapter 1

Automating Labor: Evidence from Firm-level Patent Data

This chapter is joint work with Antoine Dechelezprêtre, David Hémous, and Morten Olsen.*

1.1 Introduction

Do higher wages lead to more labor-saving innovations? And if so, by how much? At a time of fast technological progress in automation technologies such as robotics and AI and of political campaigns pushing for higher minimum wages, answering these questions is of central importance. Even more so because the endogeneity of automation innovations matters for the long-term effects of policy interventions (Hémous and Olsen, 2018). Yet, the literature on the effect of wages on labor-saving

*Antoine Dechezleprêtre, Grantham Research Institute and Centre for Economic Performance, LSE and OECD, antoine.dechezlepretre@oecd.org. David Hémous, University of Zurich and CEPR, david.hemous@econ.uzh.ch, Morten Olsen, University of Copenhagen, mgo@econ.ku.dk. Carlo Zanella, University of Zurich, carlo.zanella@econ.uzh.ch. David Hémous gratefully acknowledges the financial support of the European Commission under the ERC Starting Grant 805007 AUTOMATION. We thank Lorenzo Casaburi, Nir Jaimovich, Elias Papaioannou, Pascual Restrepo, Joachim Voth, and Fabrizio Zilliboti among others for helpful comments and suggestions. We also thank seminar and conference participants at the University of Zurich, Swiss Macro workshop, the University of Copenhagen, the TRISTAN workshop in Bayreuth, the University of Bath, London Business School, the NBER Macroeconomics Across Time and Space Conference, the NBER Summer Institute, the AlpMacro Conference, LMU, Oxford University and Helsinki Graduate School of Economics. We thank Amedeo Andriollo, Selina Schön, and Shi Suo for fantastic research assistance.

technological change remains limited. In fact, the few existing papers focus on the effect of labor costs on the *adoption* of automation technologies (e.g. Lewis, 2011, Hornbeck and Naidu, 2014, or Acemoglu and Restrepo, 2018a). Our paper is the first to establish a causal effect of an increase in wages on automation *innovations*.

Answering this question requires overcoming two challenges: identifying automation innovations and finding a source of exogenous variation in wages from the perspective of innovating firms. To overcome the first challenge, we build a new method of classifying automation patents using the existing assignment of patents to technological categories (IPC and CPC codes). We use the text of patents from the European Patent Office (EPO) and compute the frequency of certain keywords (such as "robot", "automation" or "computer numerical control") for each technological category. Our identification strategy is ideally suited for innovations in equipment and we restrict attention to those innovations. We define "automation technological categories" as technological categories where the frequency of the keywords is above a certain threshold. Finally, we identify as automation patents those which belong to automation technological categories (including non-EPO patents). Our method presents at least two advantages: it is transparent and covers a wide range of innovations across several sectors compared with more narrow measures such as the use of robots. According to our laxer measure, the share of automation innovations among innovations in machinery has increased from 12.8% in 1999 to 20.5% in 2014. We conduct a validation exercise based on Autor, Levy and Murnane (2003). We find that in the United States, sectors where the share of automation patents filed in machinery was high, saw a decrease in routine tasks and an increase in the skill ratio. Automation is uncorrelated with computerization and captures a different form of technological change but has similar effects.

At the country level, technology and wages are co-determined. To isolate exogenous variation in wages, we therefore exploit firm-level variations in the wages faced by the potential customers of innovating firms by exploiting variations in innovating firms' exposure to international markets. We expand on the methodology of Aghion, Dechezleprêtre, Hémous, Martin and Van Reenen (2016, henceforth ADHMV) and use the PATSTAT database, which contains close to the universe of patents. For each firm which undertakes automation innovations, we compute how much it has patented pre-sample in machinery in each country. We take this information as a proxy for the distribution of the firm's international exposure and build firm-specific weighted averages of low- and high-skill wages using country-level data. These firm-specific wages proxy for the average wage paid by the downstream firms of the innovating firms. As a result, for, say, two German firms, we identify the effect of an increase in wages on automation innovations, by comparing how much more automation innovations increase for the firm which has the higher market exposure to the US when US low-skill wages increase.

We conduct our main analysis over the sample period 1997-2011 and use wage data for 41 countries with automation patents for 3,341 firms. We find a substantial effect of wages on automation innovations: higher low-skill wages lead to more automation innovations with an elasticity between 2 and 4 depending on specification. Higher high-skill wages tend to reduce automation, but with a smaller elasticity, a finding in line with the capital-skill complementarity hypothesis (Krusell, Ohanian, Rios-Rull and Violante, 2000). Our results are robust to the inclusion of domestic country-year fixed effects and continues to hold when we decompose firm-specific wages into a domestic and a foreign part. Moreover, we use the geographical localization of firms' inventors to compute the local knowledge stocks which firms are exposed to. We find strong evidence of local knowledge spillovers which suggest that the long-term effects of an increase in wages on automation innovations are larger than the short-term effects. Yet, more automation innovations in a firm are associated with fewer future automation innovations. We run placebo regressions with low-automation patents in machinery and find no effect of low-skill wages on automation innovation.

Finally, we look at the effect of the Hartz reform in Germany in 2002-2004, which aimed at increasing labor market flexibility making the use of labor more attractive and consequently lowering the incentive for automation innovation. We focus on patents from the countries with the highest exposure to Germany, excluding Germany itself. While foreign firms most exposed to Germany were increasingly doing automation innovations relative to other innovations in machinery until the Hartz reform, the trend sharply reversed thereafter.

The theoretical argument that higher wages should lead to more labor-saving technology adoption or innovation dates back to Habakkuk (1962) and is at the core of several theoretical papers (e.g. Zeira, 1998, Acemoglu, 2010). More recently, a small growth literature has emerged where endogenous innovation can take the form of either automation or the creation of new tasks, in which case wages affect the direction of innovation (Hémous and Olsen, 2018, Acemoglu and Restrepo, 2018b).

There is an extensive literature on the effects of technological change on wages and employment,[1] yet the empirical literature on the reverse question is much more limited. A few papers show that labor market conditions affect labor-saving technology adoption in health care (Acemoglu and Finkelstein, 2008), agriculture (Manuelli and Seshardi, 2014, Hornbeck and Naidu, 2014, and Clemens, Lewis and Postel, 2018), and manufacturing (Lewis, 2011). Lordan and Neumark (2018) find that minimum wage hikes displace workers in automatable jobs and Acemoglu and Restrepo (2018a) relate demographic trends to robot adoption. Our paper differs in at least two ways. First, our analysis is broader since it covers a range of automation technologies and 40 countries. Second, we focus on innovation instead of adoption,[2] which matters because the economic drivers of innovation may differ from those of adoption: it may be less responsive to macroeconomic variables such as wages and knowledge spillovers are likely to play a greater role. There is essentially no empirical literature on automation innovations: Alesina, Battisti and Zeira (2018) find in cross-country regressions that labor market regulations are positively correlated with innovation in low-skill intensive sectors, which is consistent with a model where innovation is low-skill labor-saving; and a recent working paper by Bena and Simintzi (2019) shows that firms with a better access to the Chinese labor market decrease their share of process innovations after the 1999 U.S.-China trade agreement.[3]

[1]See for instance Autor, Katz and Krueger (1998), Autor et. al. (2003), Bartel, Ichniowski and Shaw (2007) or Autor and Dorn (2013), Gaggl and Wright (2017) for IT, Doms, Dunne and Totske (1997) for factory automation, Graetz and Michaels (2017) or Acemoglu and Restrepo (2017) for robots, Blanas, Gancia and Lee (2018) for different forms of capital, Mann and Püttmann (2018) or Bessen, Goos, Salomons and van den Berge (2019) for broader measures of automation and Aghion, Jones and Jones (2017), Martinez (2018) or Gaggl and Eden (2018) for the effect on factor shares (see also Aghion, Bergeaud, Boppart, Klenow and Li, 2019, and Akcigit and Ates, 2019, for other factors behind the drop of the labor share).

[2]To be more precise, Acemoglu and Restrepo (2018a) also show some cross-country correlations between demographic trends and patents in robotics.

[3]Process innovations and automation innovations are not the same: certain process innovations

This is perhaps surprising because a large literature shows that the direction of innovation is endogenous in other contexts: Acemoglu and Linn (2004) in the pharmaceutical industry; Hanlon (2015) in the 19th century cotton industry and several papers in the context of energy-saving or green innovations (Newell, Jaffe and Stavins, 1999, Popp, 2002 and Calel and Dechezleprêtre, 2016). Here, we build more specifically on the methodology of ADHMV, who build firm-level variations in gas prices to show that higher gas prices lead firms in the auto industry to engage more in clean and less in dirty innovations.[4]

The use of text analysis using keywords has developed rapidly in economics since Gentzkow and Shapiro (2010). More closely related, Mann and Püttman (2018) use machine-learning techniques to classify automation patents. We compare our approaches below.

Section 1.2 contains our first contribution: a classification of automation technologies and compares it with existing measures. Section 1.3 introduces a simple model to motivate the analysis. Section 1.4 describes our empirical strategy and the data we use. Section 1.5 contains the results of the main analysis on the effect of wages on automation innovations. Section 1.6 discusses the event study of the Hartz reform. Section 1.7 concludes. Appendix 1.8 provides details on our automation classification and additional robustness checks.

1.2 Classifying automation patents

In the following we describe the patent data as well as our method for classifying automation patents. We then show how our measure of automation compares to previous measures of automation, notably the use of computers in the framework of Autor et. al. (2003). Our approach proceeds in three steps: i) We use the

may involve reducing other costs than labor costs (for instance materials costs) and certain automation innovations can be product innovations (for instance a new industrial robot is a product innovation for its maker).

[4]Three other papers have used ADHMV's methodology: Noailly and Smeets (2015) use it to look at innovation in electricity generation, Coelli, Moxnes and Ulltveit-Moe (2018) use it to look at the effect of trade policy on innovation and Aghion, Bénabou, Martin and Roulet (2019) to look at the role of environmental preferences and competition in innovation in the auto industry—as explained later in the text, we methodologically extend this work by looking separately at the effect of the domestic and foreign variables.

existing literature to identify keywords related to automation. ii) We use those keywords and the text of EPO patents to classify technological categories (based on the existing IPC and CPC codes) in machinery as automation or not. iii) We then classify worldwide patents as automation or not depending on whether they belong to an automation technology category.

1.2.1 Patent data

We use two patent databases maintained by the European Patent Office (EPO). For most of our empirical analysis, we use the World Patent Statistical Database (PATSTAT) from Autumn 2018 which contains the bibliographical information of patents from 90 patent-issuing authorities (covering nearly all patents in the world) but not the text of individual patents. Since text analysis is essential to our approach, we supplement with the EP full-text database from 2018, which contains the full text of EPO patent applications (a subset of the patents from PATSTAT).

PATSTAT allows us to identify "patent families", a set of patent applications across different patent offices which represent the same innovation. For each patent family, we know the date of first application (which we use as the year of an innovation), the patent offices where the patent is applied for (which indicates its geographical breadth), the identity of the applicants and the inventors and the number of citations received by the patent family. In addition, to identify the technological characteristics of patents we use their IPC and their CPC codes (henceforth C/IPC codes).[5] Importantly each patent usually has several C/IPC codes. The C/IPC codes form a hierarchical classification systems. Certain types of technologies (for instance fossil fuel engines) can readily be identified to existing groupings of C/IPC codes. Such a grouping does not exist for automation and it is our goal in the following to create it.

Our strategy to identify automation innovations relies on first identifying automation C/IPC codes (and combinations thereof) by computing the frequency of certain keywords in the text of patents belonging to those C/IPC codes. We then use

[5]The IPC is the International Patent Classification and the CPC the Cooperative Patent Classification used by the USPTO and the EPO. The CPC is an extension of the IPC and contains around 250,000 codes in its most disaggregated form.

this information to identify automation patents as those with automation C/IPC codes. This strategy has two advantages over classifying patents directly. First, it allows us to include non-EPO patents in our analysis, for which PATSTAT does not contain the text.[6] Second, technological codes by themselves are informative and one should think of the particular wording of a patent as a signal of its underlying characteristics. Patents are written in different styles, and often do not expand on the purpose of the invention, so that the same innovation can often be described with or without using our keywords. In other words, if a patent does not contain one of our keywords but belongs to a C/IPC code where patents most of the time do, there is a high likelihood that it is actually an automation patent (see examples in Figures 1.2a and 1.2b below). Conversely, if a patent uses one of our keywords but does not belong to any C/IPC codes where this is common, the inclusion of this keyword is frequently uninformative about the nature of the innovation.[7]

1.2.2 Choosing automation keywords

In the following we explain how we choose our automation keywords. Most of our keywords come from the automation technologies identified in Doms, Dunne and Troske (DDT, 1997) and Acemoglu and Restrepo (AR, 2018).[8] We complemented this list as described below. Naturally, we seek to capture as many patents truly associated with automation as possible without too many false positives. Table 1.1 describes the list of keywords together with their origin (Appendix 1.8.2 provides additional details).

We have eight categories of keywords. Five of these, Robot*, numerical control, computer-aided design and manufacturing, flexible manufacturing and pro-

[6] To give an idea of the increase in sample, over the period 1997-2011 there are 3.19 million patent families with patent applications in at least two offices (a condition we will impose in our main analysis). Among those only around 740 thousand have an EPO patent with a description in English.

[7] As a matter of fact, the World Intellectual Property Organization (WIPO) offers on its website a simple tool based on a similar principle. A search engine allows to identify up to 5 IPC codes most likely to correspond to a set of keywords using the text of the patents in its database.

[8] Doms, Dunne and Troske (1997) measure automation using the Survey of Manufacturing Technology (SMT) from 1988 and 1993 conducted by the US Census. The survey asked firms about their use of certain automation and information technologies. Acemoglu and Restrepo (2018) include imports of automation technology and associate specific HS-categories from Comtrade with automation technology.

Table 1.1: Choice of automation keywords

Key words	Comments	Source
Automat*	*Automation, automatization* *or automat* at least 5 times* or (automat* or autonomous) with (secondary words or *warehouse or operator* or arm or convey* or handling or inspect or knitting or manipulat* or regulat* or sensor or storage or store or vehicle system or weaving or welding) in the same sentence at least twice	Own / Doms, Dunne and Troske (DDT) / Acemoglu and Restrepo (AR)
Robot*	Not surgical or medical	DDT and AR
Numerical Control	CNC or numeric* control* or (NC in the same sentence as secondary words)	DDT and AR
Computer-aided design and manufacturing	Computer-aided/-assisted/-supported In the same patent as secondary words CAD or CAM in same sentence as secondary words	DDT
Flexible manufacturing		DDT
Programmable logic control	Programmable logic control or PLC and not (powerline or "power line")	DDT
3D printer	*Including additive layer manufacturing*	Own
Labor	*Including laborious*	Own
Secondary words	*Machine or manufacturing or equipment or apparatus or* machining	

Notes: "In the same sentence as control words" refers to at least one control word. Keywords include i) natural adjacent words (i.e. numerical control includes NC, numerically controlled and numeric control), ii) British/American spelling (i.e. labour/labor) and iii) hyphenated adjectives (i.e. computer aided / computer-aided design). We added words in italics, the others come from AR or DDT. See Appendix for details.

grammable logic control are automation technologies in DDT or AR. Simply applying these keywords may result in false positives. For instance "NC" can refer to either "numeric control" or "North Carolina". To address this issue, we require that those keywords are either in the same patent or the same sentence as a list of secondary words which indicate that the text describes a machine. We add 3D printing, which was in its infancy when DDT was written. We also add "labor" which indicates that an innovation reduces labor costs.

We similarly add "automation" and "automatization". The stem "automat*" gather too many false positives such as "automatic transmission". We resolve this in two ways: either by restricting attention to patents where the frequency is 5 or more or by combining automat* with other words which largely come from technologies described in DDT or AR (we count patents where automat* and one of these words appear in the same sentence at least twice).

An alternative procedure would have been to read and classify a subset of patents and use machine-learning techniques to classify patents (or technological categories) as automation or not. This is the procedure in Mann and Püttmann (2018). We believe our approach has several advantages. First, we found that classifying patents as automation is a difficult task: often looking at a single patent in isolation is not enough, and one needs to look at several patents within the same technological group to find patterns suggesting that a patent is likely an automation patent. Therefore, the task of manually classifying patents cannot be easily systematized and therefore outsourced. Second, patents are written in a technical language and do not primarily discuss the goal of an innovation, so that only a few words within the text are informative. Consequently, a machine-learning algorithm would require a large set of classified data to classify patents correctly. Third, once the classification is done it can easily be applied to patents without text and future patents. Fourth, our method is more transparent and can easily be replicated or modified.

1.2.3 Defining automation technological codes and automation patents

As discussed above we use the keywords to associate technological categories, and not patents directly, to automation. These technological categories are defined as:

6-digit C/IPC codes, all pairs of 4-digit C/IPC codes and pairs combining the union of the 3 digit codes G05 and G06 with any 4-digit C/IPC codes (outside codes in G05, G06).[9] The code G05 corresponds to "controlling; regulating" and G06 to "computing; calculating; counting". Using combinations of G05 and G06 code with 4-digit C/IPC codes is inspired by Aschhoff et al. (2010) who use these codes to identify advanced manufacturing technologies. We restrict attention to categories which contain at least 100 patents to ensure that the prevalence of keywords measure is based on a sufficiently large number of patents.[10]

We then measure the prevalence of our keywords within technological categories for those patent applications from 1978 onward which contain a description in English (a total of 1,538,370 patent applications). In Appendix 1.8.2, we verify that the choice of the starting year does not affect our classification much. To select automation C/IPC codes, we further restrict attention to C/IPC codes which belong to technological fields which are associated with equipment. There are 34 technological fields (see Figure 1.8.7) and we focus on "machine tools", "handling", "textile and paper machines" and "other special machines", which we refer to as "machinery" patents (we use machinery and equipment interchangeably). Our classification scheme captures a broader set of automation technologies than what is relevant for our empirical analysis including Roombas and military drones. We adjust the set of technological codes accordingly.[11] For pairs of 4 digit IPC codes, we assume that they belong to the relevant technological field when at least one of the 4 digit codes

[9]Technically, the structure of the C/IPC classification is as follows: C/IPC "classes" have 3 digit codes (for instance B25: "hand tools; portable power-driven tools; handles for hand implements; workshop equipment and manipulators"), "subclasses" have 4 digit codes (for instance B25J: "manipulators; chambers provided with manipulation devices") and main groups have 5 to 7 digit codes (for instance B25J 9: "programme-controlled manipulators"). In the following, we will slightly abuse language and refer to classes, subclasses and main groups as 3 digit, 4 digit and 6 digit codes respectively.

[10]We group 6-digit codes with less than 100 patents into codes at the 4-digit level.

[11]Roombas are already excluded since they are not in the four technological fields. We further exclude F41 and F42 which correspond to weapons and ammunition and are in "other special machines". In addition, we include B42C which corresponds to machines for book production and B07C which corresponds to machines for postal sorting as both correspond to equipment technologies and contain 6-digit codes with a high prevalence of automation keywords; the 6-digit code G05B19 which corresponds to "programme-control systems" and contains a large number of NC and CNC (computer numerically controlled) machine tools which are not attributed IPC codes in the machine tools technological field; and the 6-digit code B62D65 which concerns engine manufacturing even though the rest of the B62D code is about the vehicle parts themselves.

belongs to the relevant technological field. Similarly, the combinations of 4 digit IPC code and G05 or G06 belong to the relevant technological fields if the 4 digit code belongs to that group.

We extensively checked the C/IPC codes and sampled patents from each category to ensure that the procedure delivered reasonable results. However, the validation exercises and the main empirical exercise where carried out after the classification was set.

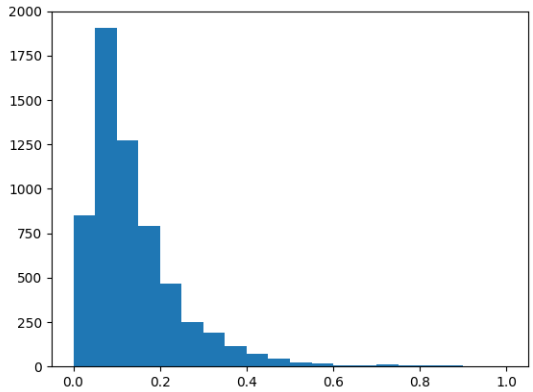
Table 1.2 gives some examples of 6-digit C/IPC codes in machinery with the prevalence of automation keywords including their rank within machinery 6 digit codes with at least 100 patents. It also shows the prevalence of some of the most important subcategories (automat*, robots and CNC) in the patents linked to each C/IPC code. C/IPC codes associated with robotics (B25J) have the highest prevalence numbers with up to 91% patents in B25J5 which contain at least one of the keywords. Yet, there are also codes associated with machine tools other than robots at the top of the distribution such as B23Q15 and codes associated with devices used in the agricultural sector such as A01J7. B24B49 is a code close to the threshold we use to delimit automation patents. The last four C/IPC codes are examples with a low prevalence of automation keywords. The table also shows that the different sub-measures do not capture the same technologies: the robotic codes are ranked highly thanks to their share of patents with the word "robot", B23Q15 is high because a lot of patents contain words related to CNC, and B65G1, because a lot of patents contain words associated with automation directly.

Figure 1.1 gives the histograms of the prevalence of automation keywords for all C/IPC 6 digit codes (panel a) and C/IPC 6 digit codes in the "machinery" technological field (panel b). The histograms show that most C/IPC codes have a low prevalence of automation keywords and that the distribution is shifted to the right for the relevant technological fields. Yet, a few codes have a high prevalence measure. Appendix 1.8.2 gives additional statistics on the prevalence measures.

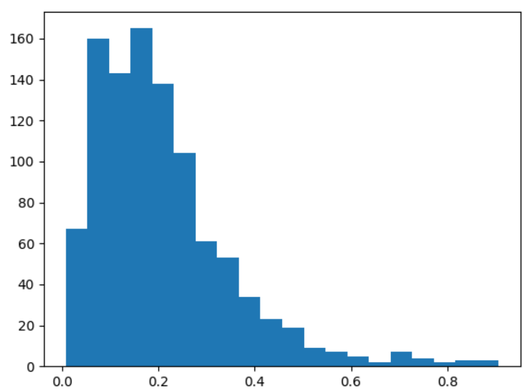
Consequently, we define automation technological categories as those with a prevalence measure above some threshold. As our baseline, we choose thresholds at the 90^{th} and 95^{th} percentiles of the 6 digit code distribution within the machinery technological field, which are given by 0.386 and 0.477 respectively. We then define a

Code	Description	Number of patents	All share	Rank (over 1009)	Robot share	Automat* share	CNC share
		– High prevalence –					
B25J5	Manipulators mounted on wheels or on carriages.	504	0.91	1	0.87	0.27	0.01
B25J19	Accessories fitted to manipulators, e.g. for monitoring or for viewing; safety devices combined with or specially adapted for use in connection with manipulators.	1001	0.89	2	0.85	0.22	0.04
B25J13	Controls for manipulators.	857	0.88	3	0.81	0.27	0.03
B25J9	Programme-controlled manipulators.	2809	0.86	4	0.79	0.29	0.07
B23Q15	Automatic control or regulation of feed movement, cutting velocity or position of tool or work.	591	0.79	7	0.09	0.36	0.65
A01J7	Accessories for milking machines or devices.	395	0.77	9	0.62	0.52	0
G05B19	Programme-control systems.	7133	0.70	16	0.22	0.39	0.25
B65G1	Storing articles, individually or in orderly arrangement, in warehouses or magazines.	1064	0.58	29	0.18	0.46	0.01
B24B49	Measuring or gauging equipment for controlling the feed movement of the grinding tool or work; Arrangements of indicating or measuring equipment, e.g. for indicating the start of the grinding operation.	608	0.42	75	0.12	0.18	0.19
		– Low prevalence –					
B65H7	Controlling article feeding, separating, pile-advancing, or associated apparatus, to take account of incorrect feeding, absence of articles, or presence of faulty articles.	736	0.28	228	0.01	0.25	0.00
B23P6	Restoring or reconditioning objects.	613	0.26	266	0.07	0.06	0.05
A01B63	Lifting or adjusting devices or arrangements for agricultural machines or implements.	264	0.24	306	0.01	0.20	0
B66D3	Portable or mobile lifting or hauling appliances.	215	0.13	677	0.02	0.07	0.00

Table 1.2: Examples of 6-digit C/IPC codes in relevant technological fields



(a) For all C/IPC 6 digit codes



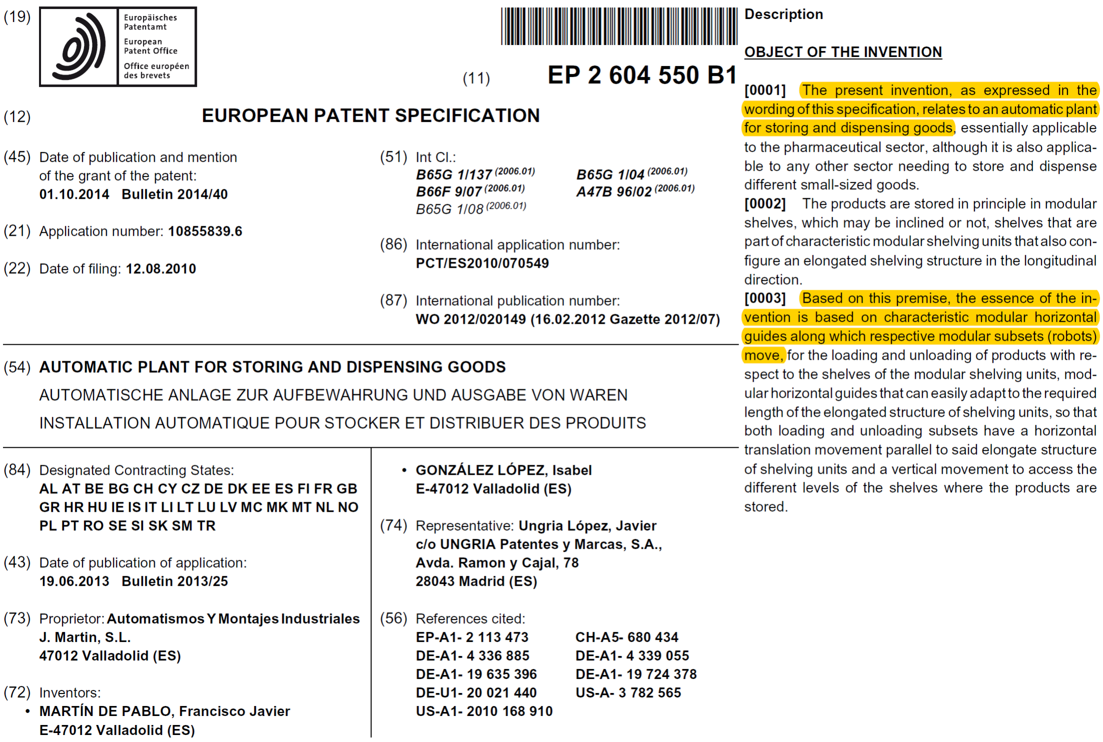
(b) For C/IPC 6 digit codes in machinery with at least 100 patents

Figure 1.1: Histogram of the prevalence of automation keywords for C/IPC 6 digit codes

patent as an automation patent if it belongs to at least one automation technological group (that is a 6 digit code, a pair of 4 digit codes, or a combination of 4 digit code and G05/G06).[12] We refer to the two classifications as auto90 and auto95 depending on the threshold used. We can similarly define subcategories of automation patents such as robot90 which correspond to patents which contain at least one technological group for which the frequency of the keywords related to robots (uniquely) is above the threshold defining auto90. By definition all robot90 patents are also auto90 patents.

Figure 1.2 shows two automation patents. Both are automated storage cabinets and are counted as automation patents because they contain the IPC 6 digit code B65G 1. As described in Table 1.2, B65G 1 corresponds to devices for storing articles and has a high prevalence of automation keywords (0.58, which is above the 95^{th} percentile threshold). The patent of Figure 1.2a contains our keywords: a sentence with the words "automatic" and "storing," and another sentence with the word "robot." The description strongly suggests that this is indeed an automation patent. The patent of Figure 1.2b does not contain any of the keywords, but the description of the text still describes a labor-saving innovation.

[12] In practice, most automation patents in our dataset are automation patents because they belong to at least one 6 digit automation code—see Appendix 1.8.2 for more details.



(19) Europäisches Patentamt European Patent Office Office européen des brevets

(11) EP 2 604 550 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent: **01.10.2014 Bulletin 2014/40**

(21) Application number: **10855839.6**

(22) Date of filing: **12.08.2010**

(51) Int Cl.: *B65G 1/137* (2006.01) *B65G 1/04* (2006.01) *B66F 9/07* (2006.01) *A47B 96/02* (2006.01) *B65G 1/08* (2006.01)

(86) International application number: **PCT/ES2010/070549**

(87) International publication number: **WO 2012/020149 (16.02.2012 Gazette 2012/07)**

(54) **AUTOMATIC PLANT FOR STORING AND DISPENSING GOODS**
AUTOMATISCHE ANLAGE ZUR AUFBEWAHRUNG UND AUSGABE VON WAREN
INSTALLATION AUTOMATIQUE POUR STOCKER ET DISTRIBUER DES PRODUITS

(84) Designated Contracting States: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK SM TR**

(43) Date of publication of application: **19.06.2013 Bulletin 2013/25**

(73) Proprietor: **Automatismos Y Montajes Industriales J. Martin, S.L. 47012 Valladolid (ES)**

(72) Inventors:
- **MARTÍN DE PABLO, Francisco Javier E-47012 Valladolid (ES)**
- **GONZÁLEZ LÓPEZ, Isabel E-47012 Valladolid (ES)**

(74) Representative: **Ungria López, Javier c/o UNGRIA Patentes y Marcas, S.A., Avda. Ramon y Cajal, 78 28043 Madrid (ES)**

(56) References cited:
EP-A1- 2 113 473 CH-A5- 680 434 DE-A1- 4 336 885 DE-A1- 4 339 055 DE-A1- 19 635 396 DE-A1- 19 724 378 DE-U1- 20 021 440 US-A- 3 782 565 US-A1- 2010 168 910

Description

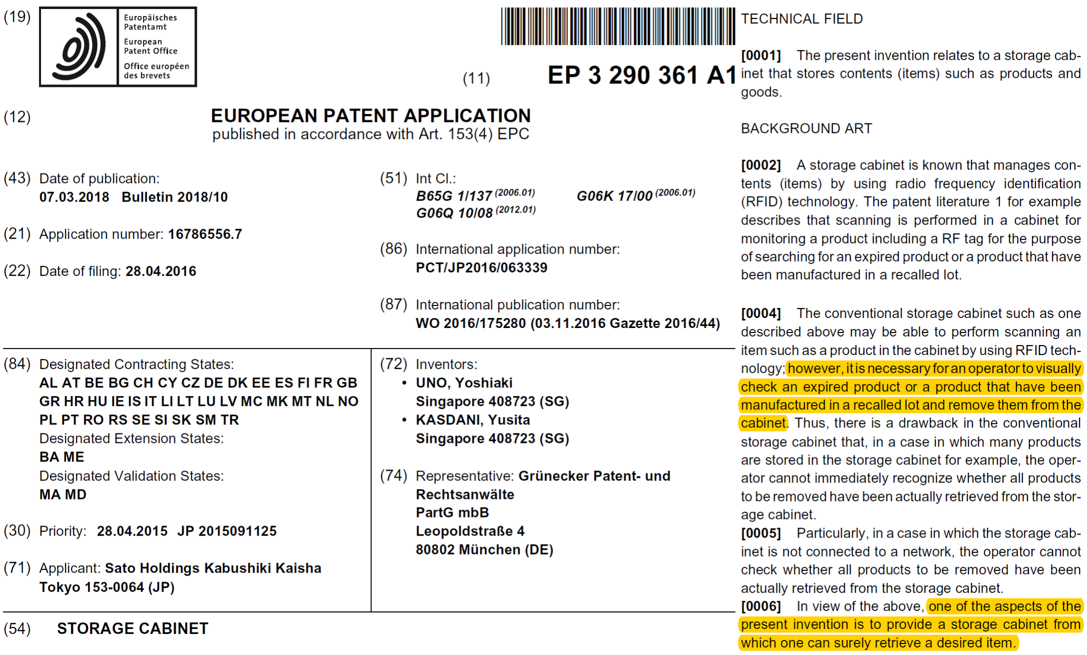
OBJECT OF THE INVENTION

[0001] The present invention, as expressed in the wording of this specification, relates to an automatic plant for storing and dispensing goods, essentially applicable to the pharmaceutical sector, although it is also applicable to any other sector needing to store and dispense different small-sized goods.

[0002] The products are stored in principle in modular shelves, which may be inclined or not, shelves that are part of characteristic modular shelving units that also configure an elongated shelving structure in the longitudinal direction.

[0003] Based on this premise, the essence of the invention is based on characteristic modular horizontal guides along which respective modular subsets (robots) move, for the loading and unloading of products with respect to the shelves of the modular shelving units, modular horizontal guides that can easily adapt to the required length of the elongated structure of shelving units, so that both loading and unloading subsets have a horizontal translation movement parallel to said elongate structure of shelving units and a vertical movement to access the different levels of the shelves where the products are stored.

(a) Example with keywords



(19) Europäisches Patentamt European Patent Office Office européen des brevets

(11) EP 3 290 361 A1

(12) EUROPEAN PATENT APPLICATION
published in accordance with Art. 153(4) EPC

(43) Date of publication: **07.03.2018 Bulletin 2018/10**

(21) Application number: **16786556.7**

(22) Date of filing: **28.04.2016**

(51) Int Cl.: *B65G 1/137* (2006.01) *G06K 17/00* (2006.01) *G06Q 10/08* (2012.01)

(86) International application number: **PCT/JP2016/063339**

(87) International publication number: **WO 2016/175280 (03.11.2016 Gazette 2016/44)**

(84) Designated Contracting States: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
Designated Extension States: **BA ME**
Designated Validation States: **MA MD**

(30) Priority: **28.04.2015 JP 2015091125**

(71) Applicant: **Sato Holdings Kabushiki Kaisha Tokyo 153-0064 (JP)**

(72) Inventors:
- **UNO, Yoshiaki Singapore 408723 (SG)**
- **KASDANI, Yusita Singapore 408723 (SG)**

(74) Representative: **Grünecker Patent- und Rechtsanwälte PartG mbB Leopoldstraße 4 80802 München (DE)**

(54) **STORAGE CABINET**

TECHNICAL FIELD

[0001] The present invention relates to a storage cabinet that stores contents (items) such as products and goods.

BACKGROUND ART

[0002] A storage cabinet is known that manages contents (items) by using radio frequency identification (RFID) technology. The patent literature 1 for example describes that scanning is performed in a cabinet for monitoring a product including a RF tag for the purpose of searching for an expired product or a product that have been manufactured in a recalled lot.

[0004] The conventional storage cabinet such as one described above may be able to perform scanning an item such as a product in the cabinet by using RFID technology; however, it is necessary for an operator to visually check an expired product or a product that have been manufactured in a recalled lot and remove them from the cabinet. Thus, there is a drawback in the conventional storage cabinet that, in a case in which many products are stored in the storage cabinet for example, the operator cannot immediately recognize whether all products to be removed have been actually retrieved from the storage cabinet.

[0005] Particularly, in a case in which the storage cabinet is not connected to a network, the operator cannot check whether all products to be removed have been actually retrieved from the storage cabinet.

[0006] In view of the above, one of the aspects of the present invention is to provide a storage cabinet from which one can surely retrieve a desired item.

(b) Example without keywords

Figure 1.2: Examples of automation patents from technological code B65G1, which are both automated storage cabinets.

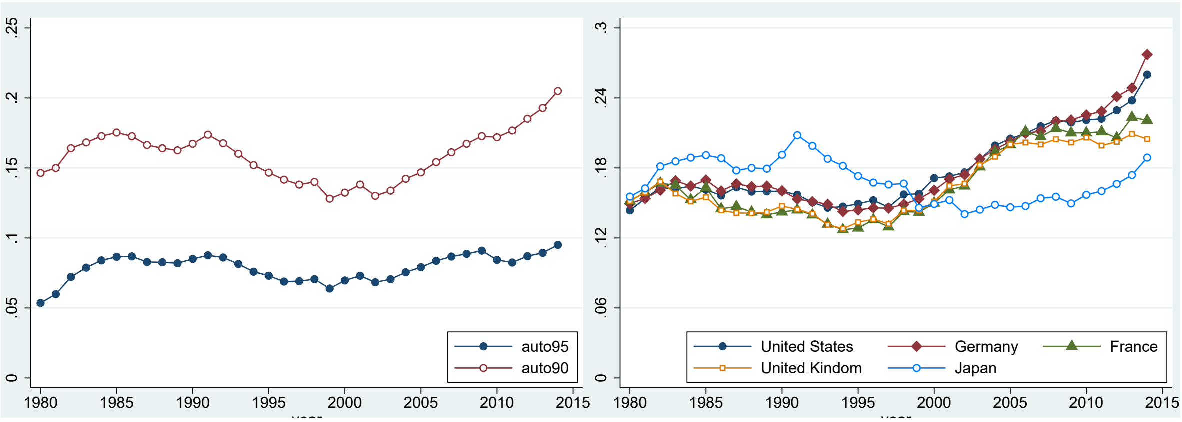
1.2.4 Trends in automation innovations

To ensure that we only capture innovations of a sufficiently high quality, we restrict attention to patent families with patent applications in at least two countries in our main empirical analysis and for the trends depicted here. We refer to these as biadic patents.[13] Several studies have documented that biadic patents are of higher quality and fundamentally different from patents applied for in only one office (e.g. Harhoff, Scherer and Vopel, 2003, van Pottelsberghe de la Potterie and van Zeebroeck, 2008, De Rassenfosse, Dernis, Guellec, Picci and van Pottelsberghe de la Potterie, 2013, and Dechezleprêtre, Ménière and Mohnen, 2017). In addition, patents can be more or less broad across countries: for instance the same invention may be covered by two patents in Japan but only one in the US. By focusing on biadic patents, we only count such a case as one innovation.[14]

Figure 1.3 below shows the evolution of automation patents in the set of biadic patents. Panel (a) shows that worldwide, the share of automation patents declines in the 1990s from 17.4% to 12.8% for the laxer auto90 measure and from 8.8% to 6.4% for the stricter auto95 measure before increasing quickly to reach 20.5% for auto90 and 9.5% for auto95 in 2014—Figure 1.8.8 in the Appendix shows that automation patents in machinery represent between 1.9 and 3.5% of all patents with the auto90 definition and it also reports the raw numbers of auto90 and auto95 patents. One interpretation is that globalization made cheap low-skill labor abroad available in the 1990s and contributed to a temporary decline in automation, which has since reversed. Panel (b) computes the share of automation patents for the auto95 measure for biadic patents conditional on the patent being protected in certain countries. The graphs show that for UK, French, German and US patents, the decline of the 1990s is less pronounced and the rise of the 2000s is very stark. In Japan, the decline of the 1990s is more pronounced and the recent growth more timid there. As a result while the share of automation patents was the highest in

[13] The original definition of biadic patents correspond to patents in at least 2 of the 3 main offices (EPO, USPTO and JPO). Our definition is a generalization counting all patent offices. We check that our results are robust to the original definition of biadic in section 1.5.6.

[14] We count patent applications and not granted patents because in certain patent offices, notably in Japan, a patent is only formally granted if the rights of the applicant are challenged. To restrict attention to patent families of even higher quality, we carry out robustness checks where we use patent citations, or patents applied to more than two offices.



(a) Share of automation patents in machinery worldwide. Automation technological categories are defined at the 90th percentile of the distribution of 6 digit C/IPC codes in machinery (for auto90) or the 95th percentile (auto95).

(b) Share of automation patents (auto95) in machinery conditional on the patent being protected in the designated countries.

Figure 1.3: Share of automation patents in machinery. Shares are computed for biadic patents.

Japan in the 1980s and early 1990s, it is now the lowest among these countries. In the Appendix, Figure 1.8.9 reports the share of automation patents in machinery according to the nationality of applicants, the trends are roughly similar but the share of Japanese patents remains higher (suggesting that the relative decline in the share of automation patents at the JPO is due to foreign firms). These country trends are similar with the auto90 measure.

1.2.5 Automation patents and robots

Recent papers (Graetz and Michaels, 2018, or Acemoglu and Restrepo, 2017) have used data on industrial robots from the International Federation of Robotics (IFR) to measure automation. The IFR reports stocks of robots by country and sectors based on yearly surveys of robot suppliers.

We first compare our automation measure with robotization at the country level. To measure robotization in a given country, we follow Acemoglu and Restrepo (2017) and use the stock of industrial robots in 2011 minus the stock of robot in 1997 divided by total employment in manufacturing in 1997 (employment data come from the OECD database). Table 1.3 reports the correlation across 27 countries between this measure of robotization and our measures of automation, namely the shares of auto95 and auto90 patents within machinery among biadic patents applied for in each country and computed over the years 1997-2011. The correlation is quite high with a coefficient of 0.38 for the auto95 measure. When we correlate robotization with the shares of robotic patents in machinery (robot90 and robot80) we find a somewhat larger coefficient (0.46) for robot80.

We then compare our two measures of automation at the sector level for the US and Germany. The IFR data contain consistent stocks of industrial robots for 17 sectors according to the ISIC Rev 4 classification between 1997 and 2011 for Germany and between 2004 and 2011 for the US (technically the IFR data aggregate the robot stocks at the level of US, Canada and Mexico). We compute robotization in each sector by taking the difference between the stocks in the two years and dividing by employment in the first period (still using OECD data). We allocate patents to these sectors according to their (family-level) 4-digit C/IPC codes using a concordance table provided by Lybbert and Zolas (2014), and similarly measure the share of auto95, auto90, robot90 and robot80 patents in machinery for each sector over the same time periods.[15] Appendix Table 1.8.15 reports shares of auto95 patents in machinery for patents granted at the USPTO, patents protected in Germany (i.e. granted German patents or granted EPO patents protected in Germany) in 1997-2011 and for all biadic patents across sectors. The shares of automation patents are very similar in the US, Germany and for the world. The three sectors with the highest shares for auto95 are always the automotive, "computer, electronic, optical and electrical products" and "other transport equipment" industries. In addition, Table 1.3 reports correlations across sectors for these measures in the US and in Germany. We find higher levels of correlations with coefficients of 0.60 and 0.56 for both US and German industries with the auto95 measure. When we use our method to focus specifically on robotic patents, we find correlation coefficients up to 0.74 and 0.78 for the robot80 measure.

[15] Lybbert and Zolas (2014) present several probabilistic concordance tables, which are based on matching industry descriptions with the title and the abstract of patents within an IPC code. This methodology cannot a priori distinguish between the sector of use of a patent and the industry of manufacture, we verify however on a few simple examples that within machinery, the classification seemed to assign patents to the sector of use (for instance textile machines are assigned to the textile industry not the equipment industry).

Table 1.3: Correlations between our automation measures and robot intensity

	(1) Across Countries	(2) Across US Industries	(3) Across German Industries
Share of automation patents in machinery (auto95)	0.383	0.602	0.560
Share of automation patents in machinery (auto90)	0.377	0.483	0.426
Share of robot patents in machinery (robot90)	0.365	0.682	0.546
Share of robot patents in machinery (robot80)	0.461	0.740	0.780
Number of observations	27	17	17

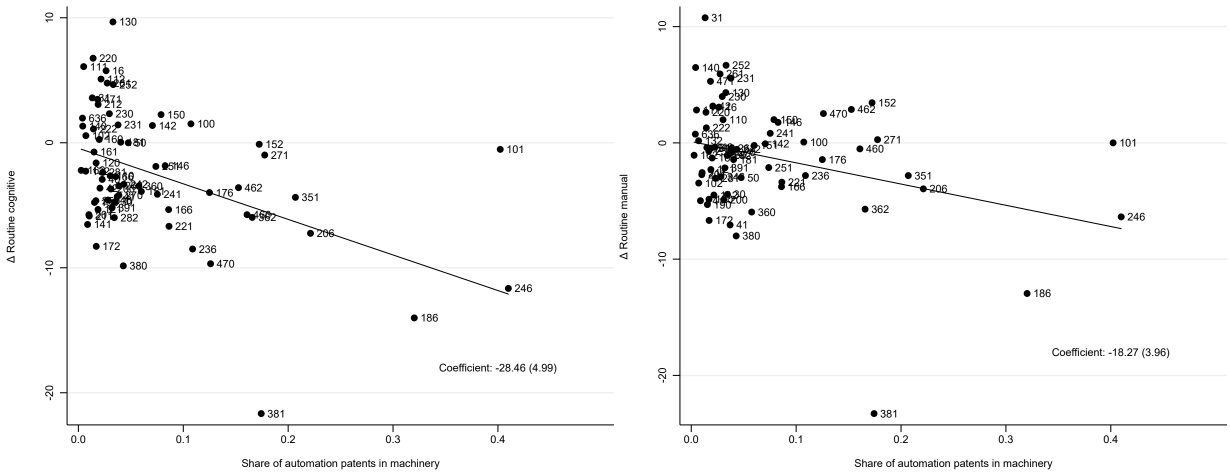
Note: This table reports correlations across countries or industries between shares of automation patents in machinery, robots patents in machinery and robot intensity. Robot intensity is measured as the difference between the stock of robots in 2011 and 1997 (columns 1 and 3) or 2004 (column 2) over employment in each country (column 1) or each sector (columns 2 and 3) in 1997 (columns 1 and 3) or 2004 (column 2). Shares of automation and robot patents are computed over the time period 1997-2011 for columns (1) and (3) and over 2004-2011 for column (2).

1.2.6 Validating our automation measure

To validate our automation measure, we use it in the framework of Autor et. al. (2003) (henceforth ALM), who show how computerization has been associated with a decrease in routine tasks at the industry level in a cross-sectional analysis on U.S. data from 1960 to 1998. Here, we provide a brief description of what we do, and we refer the reader to Appendix 1.8.3 for details. To measure automation innovations at the sectoral level, we use USPTO granted patents which belong to the machinery technological field. As before, we allocate patents to sectors according to their 4-digit C/IPC codes using a concordance table provided by Lybbert and Zolas (2014). For each sector j and each period τ, we compute the share of automation patents among machinery patents applied for during this period. We denote this variable $aut_{j\tau}$. We then run regressions of the type:

$$\Delta T_{jk\tau} = \beta_0 + \beta_C \Delta C_j + \beta_{aut} aut_{j\tau}, \tag{1.1}$$

where $\Delta T_{jk\tau}$ represents the change in tasks of type k in industry j during period τ and ΔC_j is the measure of the change of computerization in sector j (it is computed over the years 1984-1997 and used for all time periods θ). We do not first difference our measure of automation because patenting is already a measure of the flow of knowledge. We take our tasks measures directly from ALM, and therefore consider 5 types of tasks: nonroutine analytic, nonroutine interactive, routine cog-



(a) Change in routine cognitive tasks and automation intensity

(b) Change in routine manual tasks and automation intensity

Figure 1.4: Scatter plots of routine tasks changes and automation intensity (auto 95) in 1980-1998 in the United States. The list of sectors is given in Table 1.8.37

nitive, routine manual and nonroutine manual. $\Delta T_{jk\tau}$ is measured as 10 times the annual within-industry change in task input measured in percentile of the 1960 task distribution (as in ALM). We consider 3 time periods for which we can compute our automation intensity measure: 1970-1980, 1980-1990 and 1990-1998 (ALM also considers 1960-1970), and the joint time period 1980-1998. The initial concordance table mostly assigns our machinery patents to manufacturing sectors (see full list in Table 1.8.37, we restrict attention to sectors with at least 50 machinery patents per decade). As a result, we can measure automation intensity for between 67 and 69 sectors most of them in manufacturing. Our automation measures auto90 and auto95 are strongly correlated with each other (the coefficient is 0.86) but not correlated with computerization (the coefficient is 0.016 for auto95 and 0.05 for auto90).

Figure 1.4 first provides simple scatter plots of the changes in routine tasks and the share of automation patents in machinery (according to the auto95 definition) over the years 1980-1998. The list of sectors plotted (which are also the sectors in the regressions) is given in Appendix Table 1.8.37.[16] Sectors with a high share of

[16] At this level of disaggregation, the five sectors with the highest share of automation patents are: scientific and controlling instruments, optical and health services (246), dairy products (101), electronic computing equipment, office and accounting machines (186), household appliances, radio,

automation patents experience a decline in routine cognitive and routine manual tasks. Given our focus on automation in machinery a decline in routine cognitive tasks might seem surprising at first sight, but several machines replace workers for tasks such as inspection and control (such as in the example given in Figure 1.2b).

Table 1.4, columns (1) to (5) report the results of regression (1.1) for the auto95 measure. The means of the share of automation in machinery are 0.06, 0.08 and 0.07 in the 70s, 80s and 90s. Columns (3) and (4) show that sectors with a high share of automation patents in machinery experienced a large reduction in both cognitive and manual routine tasks in each decade. The coefficients of column (3) and (4) in panel B indicate that a 10 pp increase in the share of automation patents is associated with a 3 centiles and 2.2 centiles decrease in labor input of routine cognitive and manual tasks in the 1980s. To interpret a 10 pp increase, note that the standard deviation in the share of automation patents in the 1980s is 0.09, so that a 1 standard deviation increase in the automation share is associated with a decrease in routine cognitive and routine manual tasks of 2.7 and 1.9 centiles respectively. The corresponding effect of a 1 standard deviation increase in computerization is associated with a decrease in routine cognitive tasks of 0.8 centiles and essentially no change in routine manual tasks (the computerization variable has a larger effect in the 90s). We obtain similar results when we restrict attention to biadic patents (as in our main regression exercise of section 1.5) or when we exclude the equipment sector, which could be contaminated if patents are assigned to the industry of manufacture instead of the sector of use (176 in the Census classification).

Since we are interested in the effect of low- and high- skill wages on automation but do not measure the price of tasks directly, we also use the ratio of high-skill to low-skill workers (defined as college graduates over high-school dropouts and high-school graduates) as our dependent variable in cross-section regressions similar to 1.1.[17] Column (6) of Table 1.4 shows that sectors with a higher automation share also experienced a large increase in the ratio of high-skill to low-skill workers. Panel B, for instance suggests that a 10 pp increase in the share of automation patents is

TV & communications equipment, electric machinery, equipment & supplies, n.e.c., not specified electrical machinery, equipment & supplies (206) and transport equipment (351).

[17]The results are similar for the ratio of college graduates over high-school dropouts or college graduates and some college over high school graduates and dropouts.

Table 1.4: Correlation between changes in task intensity or skill ratio across sectors and automation (auto95)

	(1) Δ Nonroutine analytic	(2) Δ Nonroutine interactive	(3) Δ Routine cognitive	(4) Δ Routine manual	(5) Δ Nonroutine manual	(6) Δ H/L
Panel A: 1970 - 80, n=67						
Share of automation patents in machinery	-1.29	5.42	-17.27***	-11.43**	-1.15	0.27***
	(5.10)	(6.27)	(6.59)	(5.59)	(7.46)	(0.07)
Δ Computer use 1984 - 1997	-6.86	-3.13	-19.51***	-3.46	14.87*	0.07
	(5.72)	(7.04)	(7.41)	(6.28)	(8.38)	(0.08)
Intercept	1.06	2.31**	3.07**	2.69***	-1.75	0.05***
	(0.95)	(1.17)	(1.23)	(1.04)	(1.39)	(0.01)
R^2	0.02	0.01	0.20	0.07	0.05	0.21
Weighted mean Δ	-0.05	2.17	-0.90	1.49	0.42	0.07
Panel B: 1980 - 90, n=67						
Share of automation patents in machinery	10.09	19.05**	-30.00***	-21.61***	16.78***	1.33***
	(7.14)	(8.12)	(6.76)	(5.42)	(6.04)	(0.23)
Δ Computer use 1984 - 1997	24.80**	22.21*	-13.24	-0.42	-6.49	0.29
	(10.43)	(11.85)	(9.87)	(7.91)	(8.82)	(0.33)
Intercept	-2.62	-0.65	2.15	1.20	-2.13	-0.04
	(1.70)	(1.93)	(1.61)	(1.29)	(1.44)	(0.05)
R^2	0.12	0.14	0.27	0.20	0.11	0.37
Weighted mean Δ	1.86	4.17	-2.22	-0.59	-1.74	0.11
Panel C: 1990 - 98, n=67						
Share of automation patents in machinery	11.06*	16.02*	-22.81***	-12.53**	6.66	0.77***
	(6.08)	(8.18)	(6.54)	(5.42)	(6.28)	(0.15)
Δ Computer use 1984 - 1997	26.77***	27.00**	-23.15**	-24.87***	7.48	0.66***
	(8.35)	(11.23)	(8.98)	(7.44)	(8.62)	(0.20)
Intercept	-2.36*	-1.43	1.72	2.27*	-2.40*	-0.06*
	(1.37)	(1.84)	(1.47)	(1.22)	(1.41)	(0.03)
R^2	0.19	0.15	0.25	0.23	0.03	0.41
Weighted mean Δ	2.45	3.79	-3.44	-2.36	-0.79	0.09

Standard errors are in parentheses. Colums (1) to (5) of Panels A to C each presents a separate OLS regression of ten times the annual change in industry-level task input between the endpoints of the indicated time interval (measured in centiles of the 1960 task distribution) on the share of automation patents in machinery (defined with the 95th percentile threshold) and the annual percentage point change in industry computer use during 1984 - 1997 as well as a constant. In Column (6), the dependent variable is the ratio of high-skill (college graduates) to low-skill (high-school graduates and dropouts) workers. Estimates are weighted by mean industry share of total employment in FTEs over the endpoints of the years used to form the dependent variable. * p<0.1; ** p<0.05; *** p<0.01

associated with an increase of 1.33 in the ratio of high-skill to low-skill workers in the 1980s.

In the Appendix, Table 1.8.38 reproduces the same exercise for our laxer measure (auto90) and obtains similar results. Finally, Table 1.8.39 reproduces the same analysis separately for each education category (as ALM) and shows that automation leads to a reduction of routine tasks and an increase in non-routine manual tasks for high-school graduates (but in line with column (6) of Table 1.4, a large share of the task changes at the industry level are explained by changes in educational composition - see Panel F).

Overall, these results suggest that our automation measure captures a form of skill-biased technical change, distinct from computerization and associated with a decrease in routine tasks by low-skill workers. We can therefore use it to analyze the effect of wages on automation innovation incentives.

1.3 A simple model

Before carrying out our main empirical analysis, we present a simple one-period model to clarify our argument. The model is motivated by the business structure of the largest automation innovator. In 2018, Siemens, the biggest innovator in our sample, had 31% of its work force in Germany, but only 14% of total revenue from customers based in Germany. During this year the strongest growing division of Siemens was the Digital Factory Division which provides a broad range of automation technology to manufacturers across the globe. The annual report describes how "The Digital Factory Division offers a comprehensive product portfolio and system solutions for automation technologies used in manufacturing industries, such as automation systems and software for factory automation, industrial controls and numerical control systems, motors, drives and inverters and integrated automation systems for machine tools and production machines...". The report is centrally interested in how "Changes in customer demand [for automation technology by downstream manufacturers] are strongly driven by macroeconomic cycles" and discusses a number of such drivers including changes in cost of capital and political development towards trade protectionism.[18] Siemens further directly discusses how such macroeconomic trends affect its $R\&D$ decisions.

We incorporate these business features into a model built on the task framework of Acemoglu and Autor (2011) and more precisely on the growth model of Hémous and Olsen (2018). A manufacturing good is produced with a continuum of intermediate inputs according to the Cobb-Douglas production function:

$$Y = \exp\left(\int_0^1 \ln y\left(i\right) di\right), \tag{1.2}$$

where $y(i)$ denotes the quantity of intermediate input i. The manufacturing good is the numeraire. Each intermediate input is produced competitively with high-skill labor ($h_{1,i}$ and potentially $h_{2,i}$), low-skill labor l_i and potentially machines x_i, according to the production function:

$$y_i = h_{1,i}^{1-\beta}\left(\gamma\left(i\right) l_i + \alpha\left(i\right) \nu^\nu (1-\nu)^{1-\nu} x_i^\nu h_{2,i}^{1-\nu}\right)^\beta,$$

where $\gamma(i)$ is the productivity of low-skill workers and $\alpha(i)$ is an index which takes the value 0 for non-automated intermediates and 1 for automated intermediates. ν and β are fixed share parameters in $(0,1)$. Machines are specific to the intermediate input i. If a machine is invented, it is produced monopolistically, 1 for 1 with the final good so that the monopolist charges a price $p_x(i) \geq 1$.

At the beginning of the period, for each non-automated intermediate i, there is an innovator (Siemens). The innovator manages to create a machine specific to intermediate i with probability λ if it spends $\theta\lambda^2 Y/2$ units of manufacturing good, where θ is a productivity parameter.

We solve the model in two steps, first we derive the profits realized by machine producers, second we solve for the innovation decision. Consider an automated intermediate input (that is $\alpha(i) = 1$), then the downstream intermediate input producer is indifferent between using low-skill workers or machines together with high-skill workers in production whenever:

[18] Interestingly, the report never mentions "cost of labor" as a reason for automation, but instead used a number of euphemisms such as "increase competitiveness", "enhance efficiency", "improve cost position" and "stream line production".

$$w_H^{\nu} p_x^{1-\nu} = w_L/\gamma(i).$$

As a result, the machine producer is in "Bertrand competition" with low-skill workers. Given that a machine costs 1, the machine producer will charge a price $p_x(i) = \max\left((w/\gamma(i))^{\frac{1}{1-\nu}} w_H^{-\frac{\nu}{1-\nu}}, 1\right)$, and the intermediate input producer will use low-skill workers whenever $w_L/\gamma(i) < w_H^{\nu}$ and machines otherwise. Therefore, the machine producer can charge a higher price when low-skill wages are lower it has to charge a lower price when high-skill wages are higher since high-skill workers and machines are complement. Using that the manufacturing good is produced according to a Cobb-Douglas production function, we have that $p(i)y(i) = Y$ for all intermediates. Therefore, we can derive the profits of the machine producer for intermediate i as:

$$\pi_i^A = \max\left(1 - \left(\frac{\gamma(i)}{w_L}\right)^{\frac{1}{1-\nu}} w_H^{\frac{\nu}{1-\nu}}, 0\right) \nu\beta Y.$$

In turn, at the beginning of the period, the potential innovator solves $\max \lambda\pi_i^A - \theta\frac{\lambda^2}{2}Y$, which gives the equilibrium innovation rate as:

$$\lambda = \frac{\nu\beta}{\theta} \max\left(1 - \left(\frac{\gamma(i)}{w_L}\right)^{\frac{1}{1-\nu}} w_H^{\frac{\nu}{1-\nu}}, 0\right).$$

As a result, the number of automation innovations is equal to:

$$Aut = \frac{\nu\beta}{\theta}\int_0^1 (1-\alpha(i)) \max\left(\left(1 - \left(\frac{\gamma(i)}{w_L}\right)^{\frac{1}{1-\nu}} w_H^{\frac{\nu}{1-\nu}}\right), 0\right) di.$$

This expression is increasing in the low-skill wage w_L and decreasing in the high-skill wage w_H, with a smaller elasticity in absolute value. Intuitively, the incentive to replace low-skill workers with machines (and high-skill workers) increases with low-skill wages and make manufacturing firms better customers of machines and the reverse for high-skill wages. An upward shift in the low-skill workers productivity function $\gamma(i)$ also reduces the number of automation innovations.

More generally, the defining characteristic of automation is that it allows for the replacement of workers by machines in certain tasks. When intermediates have a unit-elasticity of substitution as in (1.2), the aggregate production function is Cobb-Douglas and automation corresponds to a change in factor shares. When intermediates have an elasticity of substitution lower than 1, the aggregate production function is CES and automation corresponds to a combination of labor-augmenting and capital-depleting technical changes (see Aghion et al., 2017).

1.4 Empirical Strategy and Data

1.4.1 Empirical strategy

We now take the predictions of our model to the data. As mentioned above, innovators in automation technologies are often large companies (e.g. Siemens) which sell their automation equipment internationally. Following the logic of our model, the incentives of the downstream producers to adopt automation technology is determined by wages in their local market. As a result, the decision of innovators such as Siemens to pursue automation research in the first place depends on the wages that their potential customers face in different countries.[19]

In our baseline regression, we assume that a firm's innovation in automation is given by the following Poisson specification:

$$\begin{aligned} & PAT_{Aut,i,t} \qquad (1.3) \\ = \quad & \exp\left(\begin{array}{c} \beta_{w_L} \ln w_{L,i,t-2} + \beta_{w_H} \ln w_{H,i,t-2} + \beta_X X_{i,t-2} \\ +\beta_{Ka} \ln K_{Aut,i,t-2} + \beta_{Ko} \ln K_{other,i,t-2} + \beta_{Sa} \ln SPILL_{Aut,i,t-2} \\ +\beta_{So} \ln SPILL_{other,i,t-2} + \delta_i + \delta_t \end{array}\right) + \epsilon_{i,t}. \end{aligned}$$

$PAT_{Aut,i,t}$ denotes the number of automation patents applied for by firm i in year t. $w_{L,i,t-2}$ and $w_{H,i,t-2}$ denote the average low-skill and high-skill wages faced by the customers of firm i at time $t-2$ (we explain below how we proxy for them). Section 1.3 predicts that $\beta_{w_L} > 0$: an increase in the average low-skill wage faced by

[19]If the automation innovation is internal to the firm, then the argument follows if one interprets the innovator's customers as the different downstream production sites of the same firm.

the customers of firm i leads firm i to undertake more automation innovations. It also predicts that $\beta_{w_H} < 0$ since high-skill workers are complementary to machines. $X_{i,t}$ represents a vector of additional controls (average GDP per capita, GDP gap and labor productivity). Controlling for GDP per capita or labor productivity allows us to control for changes in productivity in the country where machines are potentially sold[20] and controlling for the GDP gap allows us to capture business cycle fluctuations and changes in demand. We include this control because the literature finds that innovation in general is affected by the business cycle (see for instance Aghion, Angeletos, Banerjee, and Kalina, 2010).

$K_{Aut,i,t-2}$ and $K_{other,i,t-2}$ denote the stocks of knowledge in automation and in other technologies of firm i at time $t-2$. These knowledge stocks are computed using the perpetual inventory method.[21] $SPILL_{Aut,i,t-2}$ and $SPILL_{other,i,t-2}$ similarly denote the stocks of external knowledge (spillovers) in automation and in other technologies which firm i has access to at time $t-2$ (we explain below how these are constructed). δ_i is a firm fixed effect and δ_t is a time fixed effect. Finally, $\epsilon_{i,t}$ is an error term, which, we assume, is uncorrelated with the other right-hand side variables. The right-hand side variables are lagged by 2 years in the baseline regressions to reflect the delay between changes in R&D investments and patent applications—we investigate the role of our timing assumption in Section 1.5.4 below.

To control for firm-level fixed effects, we use several econometric techniques. Our baseline specification uses the Hausman, Hall and Griliches (1984) method, denoted HHG, which is the count data equivalent to the within-group estimator. Technically, this method is inconsistent with equation (1.3) because it requires strict exogeneity and therefore prevents the lagged dependent variable from appearing on the right-hand side (which it does through the knowledge stock $K_{Aut,i,t-2}$). Yet, the bias is small with large T, which is the case in our baseline regression (15 years). Second, we use the Blundell, Griffith and Van Reenen (1999) method, which proxies for the fixed effect by using the pre-sample average of the dependent variable.

[20]GDP per capita could also captures non-homotheticity in preferences, for instance if higher quality goods or services are less automated.

[21]To be more specific we use $\ln(1+K)$, a depreciation rate of 15% and add a dummy indicator variable for when each of knowledge stocks—automation and others—equals zero.

1.4.2 Macroeconomic data

Our macroeconomic variables come primarily from the 2013 release of the World Input Output Tables, henceforth, WIOD (Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J., 2015). The database contains information on hourly labor costs across groups of educational attainment – low-skill, middle-skill and high-skill workers – for the manufacturing sector from 1995 to 2009 as well as value added and producer price indices. The dataset contains information on 40 countries, including all 27 EU countries of 2009. We obtained similar data from the Swiss Federal Statistical Office to add Switzerland, a large source of patents, to our analysis. For our baseline regressions, we focus on labor costs in manufacturing since our analysis in section 1.2 showed that most of our patents (89% of biadic auto95 patents in 1997-2011) are associated with manufacturing, but we check that our results are robust to using labor costs in the entire economy. Although our measures cover all labor costs, we refer to those as wages from here on for simplicity. From the same dataset, we obtain measures of labor productivity (as value added divided by hours) and producer price indices (for the whole economy and manufacturing). We obtain exchange rate and GDP data from UNSTAT and compute the GDP gap to control for business cycles.[22] Appendix 1.8.5 provides additional details. All macroeconomic variables are deflated in the same way: In the baseline regression, we first deflate nominal values by the local producer price index for manufacturing (indexed to 1995), and then we convert everything into dollars using the average exchange rate for 1995 the starting year of our regressions.

In the data low-skill workers are defined as those without a high-school diploma or equivalent and high-skill workers as those with at least a college degree. Middle-skill wages and low-skill wages are very highly correlated so in practice one should interpret our low-skill wage variable as reflecting both low-skill and middle-skill (we look at middle-skill wages in section 1.5.6).[23]

The countries with the highest low-skill wages in 1995 are Belgium, Sweden and

[22]We use a HP filter with a smoothing parameter of 6.25 on $\ln(GDP)$ to get the trend, and the GDP gap is measured as the difference between $\ln(GDP)$ and its trend.

[23]For our baseline sample of firms, included in Table 1.7 below, the correlation between low-skill and middle-skill wages is 0.94 controlling for firm and year fixed effects. It is only 0.6 for low-skill and high-skill wages. See Appendix Table 1.8.26.

Table 1.5: Low-skill wages and the skill-premium in manufacturing sector for selected countries

Country	Low-skill wages (1995$)		Skill-premium (HS wages/LS wages)	
	1995	2009	1995	2009
India	0.19	0.28	4.79	4.98
Mexico	0.89	0.61	3.90	4.21
Bulgaria	1.29	0.71	3.32	2.25
USA	11.57	13.67	2.46	3.02
Belgium	29.50	41.89	1.56	1.46
Sweden	19.92	42.16	1.73	1.33
Finland	23.41	43.63	1.20	1.46

Note: Wages data, taken from the World Input Output Database. The table shows manufacturing low-skill wages (technically labor costs) deflated by (manufacturing) producer price index and converted to US dollars using average 1995 exchange rates. Skill-premium is the ratio of high-skill to low-skill wages (labor costs). The table shows the three countries with the lowest low-skill wages in 2009, the three with the highest and the United States.

Finland with \$41.9, \$42.2 and \$43.6 respectively (in 1995 dollars). The countries with the lowest high-skill wages in 2009 are India, Mexico and Bulgaria with \$0.28, \$0.61 and \$0.71, respectively. The corresponding number for the US is \$13.7. Table 1.5 summarizes these values for these seven countries. It further shows that the ratio of high-skill to low-skill wages varies considerably across countries, even among those that have relatively similar low-skill wages. The skill-premium in the United States rose from 2.46 to 3.02 during this period while it slightly declined in Belgium from 1.56 to 1.46.

1.4.3 Computing firm's market-specific wages and spillovers

Ideally, we would like to measure the wages paid by the (actual and potential) customers of automation innovators. We do not directly observe these, and we build a proxy which is a weighted average of country-level wages where the weights reflect the market exposure of innovators. We define the average low-skill wage faced by a firm's customers $w_{L,i,t}$ as

$$w_{L,i,t} \equiv \sum_c \omega_{i,c} w_{L,c,t}, \tag{1.4}$$

where $w_{L,c,t}$ is the low-skill wage in country c at time t and $\omega_{i,c}$ is the fixed weight of country c for firm i. Firms have different exposure to different markets because of trade barriers, heterogeneous tastes of customers, or various historical accidents if exporting involves sunk cost. This measure is a shift-share instrument (Bartik, 1991). Since the weights are fixed, our identification relies on how country-level shocks affect firms differently. In fact, had we observed the wages of the customers of automation innovators, those would have suffered from reverse causality, and we would have used our measure as an instrument. We discuss the recent literature on shift-share regressions in detail in Section 1.5.5.[24]

To measure the weights, and in the absence of sales data for most firms involved in automation innovations, we follow and expand on the methodology of Aghion et al. (2016, ADHMV). We use the firm's pre-sample history of patent filing as a proxy for the market exposure of firms. When a firm applies for a patent, it applies for protection in a specific jurisdiction, and it has to pay a fixed cost whenever it wants to expand the geographic coverage of a patent. Therefore, whether a firm protects its innovations in a country or not reflects its intent to sell or license its technology in that country (see e.g. Eaton and Kortum, 1996). Taking this into account, we compute for each firm, the fraction of its patents in the relevant technological field of machinery (not only automation) protected in each country c, $\tilde{\omega}_{i,c}$ during a pre-sample period.[25] We only count patents in the machinery because some of the biggest innovators in automation technologies are large firms (Sony, Siemens, etc.) which produce a wide array of products with different specialization patterns across industries. We restrict attention to patent families with at least one citation (not counting self-citations) to exclude the lowest quality patents.[26]

[24]As we keep the weights fixed we look at how wage changes in the countries where a firm already sells affect the firm's automation innovation. A different question would have been to analyze how wage changes affect a firm's decision to enter a new market, this is beyond the scope of this paper.

[25]In Europe, firms can apply both at national patent offices and at the European Patent Office (EPO). In the latter case, firms still need to pay a fee for each country in which they want their patent to be protected. We count a patent as being protected in a given European country if it is applied for either directly in the national office or through the EPO.

[26]Including all patents generally increases the weight of the country with the most patents, in line with the finding that poor quality patents tend to be protected in fewer countries. However, further increasing the threshold from 1 to more citations does not significantly change the distribution of weights.

Patenting indicates whether the firm intends to sell in that market. However, a larger market is likely to host more firms so that the market size per firm will generally not grow 1 for 1 with size of total market. To account for this we weigh each market c by $GDP_{0,c}^{0.35}$, where $GDP_{0,c}$ is the 5 year average GDP of country c at the end of the pre-sample period.[27] As a result, the weight of country c for firm i is given by:

$$\omega_{i,c} = \frac{\tilde{\omega}_{i,c} GDP_{0,c}^{0.35}}{\sum\limits_{c'} \tilde{\omega}_{i,c'} GDP_{0,c'}^{0.35}}.$$

We compute weights for the 41 countries for which we have wage data. The weights are computed over the pre-sample period 1970-1994 to ensure that they are weakly exogenous as patent location could be influenced by shocks to innovation. We use the same weights to compute firm customers' average high-skill wage, productivity or GDP per capita.

ADHMV verify that a similar method accounts well for the sales distribution of major auto manufacturers. Coelli, Moxnes and Ulltveit-Moe (2016) carry out a more systematic exercise and verify that a similar method accounts well for aggregate bilateral trade flows and firm exports across 8 country groups in a representative panel of 15,000 firms from 7 European countries (regressing patent weights on sales weights gives a coefficient of 0.89 with a s.e. of 0.008). In Appendix 1.8.4, we similarly show that our patent weights correlate well with trade flows. [28]

Patent data also reports where firms' innovators are located. Given that knowledge spillovers have a geographical component (Hall, Jaffe and Trajtenberg, 1993), we can use this information to build a measure of the stock of knowledge to which a firm is exposed. More specifically and similarly to ADHMV, we compute the stocks of automation patents and of other patents in each country. Then, for each firm, we build a weighted average of country-level knowledge stocks, where the weights correspond to the location of their innovators pre-sample in 1970-1994.[29]

[27]Here we use Eaton et al. (2011) who estimate that the elasticity of French exports to the GDP of the destination country is 1 while the elasticity of the number of French exporters is 0.65, which gives an elasticity of the average export by firm of 0.35. ADHMV use a power of 1 on GDP instead of 0.35. We use different values in robustness checks in section 1.5.6

[28]There are three differences between our weights and those of these previous papers: we use the empirically founded exponent of 0.35 on GDP, we restrict attention to cited patent families and to patents in certain technological fields.

[29]The country stocks are built using the perpetual inventory method with a depreciation rate

To link patents with their owners, we use Orbis Intellectual Property which links 40 million patents to companies available in the Orbis financial database. For companies in the same business group, R&D decisions could happen at the group, though treating a group as one agent is often too aggressive (for instance because subsidiaries may be in different sectors). Therefore, for firms within the same business group, we normalize company names by removing non-firm specific words such as country names or legal entity types from the name and then merge firms with the same normalized name. All other firms are treated as separate entities.[30]

1.4.4 Descriptive statistics

Our basic dataset consists of applicants who have applied to at least one biadic automation patent between 1997 and 2011 (included), who have at least one patent prior to 1995 which can be used to compute weights, and who are not fully domestic (i.e. we exclude firms which have only patented in one country pre-sample). For the auto95 measure this corresponds to $3,341$ firms, which are responsible for $35,803$, or 58% of the total number of innovations (patent families). For auto90, $4,905$ firms are responsible for $61,931$, also 58% of the total. Table 1.6 gives some descriptive statistics on the number of automation patents per year and the country weights for the firms in our sample. Over the period 1997-2011, the median firm in the sample has filed 2 auto95 and 3 auto90 patent applications. The distribution is very skewed and the 99^{th} percentile firm in the sample has filed 194 automation patents for auto90 and 173 for auto95. The largest country for a given firm has on average a weight of 0.47 (for auto95). To ensure that our results are not driven solely by the largest country, which we refer to as the "domestic country" of a firm, we will include in some regressions, domestic country-year fixed effects. The second largest country has on average a weight of 0.17. The three countries with the largest weights on average are the United States, Germany and Japan. Appendix Table

of 15%. We add dummy variables for when the spillover stocks are zero.

[30]For instance, Siemens S.A., Siemens Ltd. or Belgian Siemens S.A. are merged, but Primetals Technologies Germany Gmbh which belongs to the same group remains a separate entity in our regressions.

Table 1.6: Descriptive statistics for firms in our baseline regression

Variable	Auto95		Auto90			Auto95	Auto90
Automation pantents	per year	1997-2011	per year	1997-2011		weights	
Mean	0.7	11.22	0.84	13.24	Largest country	0.47	0.46
Standard deviation	3.46	48.71	4.04	56.76	Second largest	0.17	0.18
p50	0	2	0	3	US	0.21	0.21
p75	0.27	6	0.33	7	Japan	0.17	0.15
p90	1.4	19	1.6	22	Germany	0.2	0.21
p95	3	41	3.27	50	France	0.09	0.09
p99	12	173	13.73	194	UK	0.09	0.09
Number of firms	3341		4903				

1.8.16 gives the list of the ten biggest automation patenters in our sample.[31]

1.5 Main Empirical Results

We present our main results in three steps: First, our baseline regressions use the full variation of firm low-skill wages to estimate the effect of an increase in low-skill wages on automation innovations. Second, we use country-year fixed effects to isolate the contribution of foreign wages. Third, we contrast the results on automation innovations with those on other types of machinery innovations. The rest of the section contains additional results and robustness checks.

1.5.1 Baseline results

Our baseline results are contained in Table 1.7. The dependent variable is the number of biadic patents that qualify as automation when we use a threshold of the 95th percentile for 6 digit C/IPC codes (auto95). The regression is carried over the years 1997-2011 for the dependent variable and 1995-2009 for the independent variables, a constraint imposed by the availability of wage data for a large number of countries. Skill-dependent wages are measured in the manufacturing sector and we deflate by the producer price index in the same sector.

Column (1) shows that without any controls except fixed effects, a higher low-skill manufacturing wage for the customers of an innovating firm predicts more automation innovation. The estimated coefficient is an elasticity so that an increase of

[31] For instance, for Siemens the countries with the largest weights are Germany (0.37), the USA (0.12), France (0.10), Japan (0.09) and the UK (0.07).

Table 1.7: Baseline regressions: effect of wage on automation innovations (auto95)

Dependent variable	Auto95								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	2.2000***	2.8254***	1.8160**	1.9058**	1.9992**	2.2954***	2.4627***	2.4266***	3.7365***
	(0.5123)	(0.7332)	(0.7421)	(0.7729)	(0.8223)	(0.8198)	(0.8351)	(0.8658)	(0.9116)
High-skill wage		-0.9210	-0.9009	-0.9695	-0.8698	-0.2971	-1.6180**	-1.6700*	-0.4838
		(0.7082)	(0.6715)	(0.6913)	(0.7511)	(0.6802)	(0.8033)	(0.8634)	(0.7650)
Stock automation			-0.1275***	-0.1269**	-0.1270**	-0.1239**	-0.1441***	-0.1443***	-0.1504***
			(0.0495)	(0.0496)	(0.0495)	(0.0495)	(0.0509)	(0.0510)	(0.0510)
Stock other			0.6311***	0.6296***	0.6309***	0.6260***	0.6408***	0.6407***	0.6489***
			(0.0579)	(0.0581)	(0.0581)	(0.0574)	(0.0600)	(0.0600)	(0.0595)
GDP gap				0.0210	0.0214	0.0179	0.0279*	0.0278*	0.0265*
				(0.0159)	(0.0157)	(0.0157)	(0.0158)	(0.0157)	(0.0156)
Labor productivity					-0.2551			0.1285	
					(0.8644)			(0.9199)	
GDP per capita						-1.5635*			-3.3618***
						(0.8765)			(0.8917)
Spillovers automation							0.5442*	0.5478*	0.8587***
							(0.3135)	(0.3151)	(0.3213)
Spillovers other							-0.3014	-0.3089	-0.5853**
							(0.2248)	(0.2315)	(0.2303)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	50115	50115	50115	50115	50115	50115	50115	50115	50115
Firms	3341	3341	3341	3341	3341	3341	3341	3341	3341

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

10% in the low-skill wage is associated with 22% more automation patents. Column (2) introduces high-skill wages as a control. As predicted by the model, high-skill wages enter with a negative coefficient which is smaller in magnitude than the low-skill wage (though not statistically significant). Column (3) adds control for the firm's stock of knowledge: a higher stock of automation knowledge within the firm reduces the amount of automation innovation, suggesting that firms do not become more specialized in automation technologies over time. Column (4) controls for the GDP gap, automation innovations appear to be mildly pro-cyclical with a small elasticity which is only significant at the 10% level in some specifications. Columns (5) and (6) add controls for labor productivity in manufacturing and GDP per capita. Labor productivity does not have a significant effect and GDP per capita has a negative effect, though its significance is not robust to the specifications to follows. Columns (7) to (9) repeat columns (4) to (6) but include knowledge spillovers and find that firms which are exposed to more knowledge in automation technologies innovate more in automation (with an elasticity between 0.54 and 0.86 depending on specifications). In all specifications, the coefficient on low-skill wages is highly significant with elasticities between 1.8 and 2.8 for columns (1) to (8) and a larger elasticity of 3.7 in column (9).

Firms in the same country could be affected by common shocks, and we therefore we cluster standard errors at the domestic country (i.e. the country of largest weight) level in Appendix Table 1.8.17. If anything clustering at the country level tends to reduce the standard error on low-skill wages.[32]

Appendix Table 1.8.18 repeats Table 1.7 for the auto90 measure of automation. The results are very similar but the coefficients on low-skill wages tend to be of a smaller magnitude, which is in line with auto95 measure being a stricter measure of automation. This also helps explain the magnitude of our elasticities in Table 1.7: our analysis focuses on innovations with a high automation content (and therefore most likely to respond to an increase in wages) for firms which introduce at least one of those innovations.

1.5.2 Focusing on foreign wages

Country-level shocks which we have not controlled for may affect both innovation and wages. Insofar as firms are mainly affected by the shock of their domestic country, we can capture those through domestic country-year fixed effects. Country-year fixed effects would for instance control for a tax reform in Germany that would affect both the innovation incentives of Siemens and low-skill wages. It would also control for a technology shock that leads German firms to introduce more automation innovations and affect wages. Our identification assumption then becomes that foreign wages are exogenous to automation innovations given our set of controls. One remaining concern would arise from shocks to the cost of innovation if firms innovate outside of their domestic country. We address this issue directly in section 1.5.6 by including wages weighted by the location of the firm's inventors.[33] Furthermore, in section 1.5.3 we look at the effect of wages on low-automation machinery innovations and therefore any remaining bias would have to affect both types of

[32]A potential explanation for the negatively correlated error terms, is that a successful automation innovation by one firm will reduce the incentive for its competitors since the market has already been captured.

[33]A related concern arises from offshoring: the cost of machine production would then be correlated with foreign wages. Note, however, that higher foreign low-skill wages in production would increase the price of machines and therefore bias our coefficient on low-skill wages toward 0.

Table 1.8: Country-year fixed effects

Dependent variable	Auto95								
	Domestic + Foreign			Foreign					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	1.8852*	2.1429*	3.0411**	3.4891***	4.3023***	3.7989**	3.6420***	4.3362***	3.8663**
	(1.0367)	(1.1505)	(1.2232)	(1.2958)	(1.4482)	(1.6370)	(1.3146)	(1.4473)	(1.6288)
High-skill wage	-2.4820**	-1.9117*	-1.7526	-3.5161***	-2.4740*	-3.3526**	-3.7549***	-2.8325**	-3.6398***
	(1.0115)	(1.0157)	(1.1046)	(1.2515)	(1.4209)	(1.3633)	(1.2805)	(1.4364)	(1.3692)
GDP gap	0.0623*	0.0620*	0.0646*	0.0044	0.0016	0.0044	0.0031	0.0001	0.0031
	(0.0343)	(0.0342)	(0.0343)	(0.0492)	(0.0492)	(0.0492)	(0.0494)	(0.0494)	(0.0494)
Labor productivity		-1.2851			-1.7494			-1.5475	
		(1.6381)			(1.4131)			(1.3896)	
GDP per capita			-2.8260			-0.5289			-0.3829
			(2.0242)			(1.9347)			(1.8713)
Stock automation	-0.1511***	-0.1506***	-0.1541***	-0.1522***	-0.1523***	-0.1526***	-0.1530***	-0.1532***	-0.1533***
	(0.0528)	(0.0527)	(0.0523)	(0.0525)	(0.0523)	(0.0525)	(0.0524)	(0.0521)	(0.0524)
Stock other	0.6549***	0.6556***	0.6555***	0.6494***	0.6471***	0.6490***	0.6496***	0.6475***	0.6493***
	(0.0602)	(0.0602)	(0.0598)	(0.0602)	(0.0601)	(0.0600)	(0.0601)	(0.0601)	(0.0599)
Spillovers automation	1.4782***	1.4762***	1.4715***	1.4396***	1.4128***	1.4355***	1.4380***	1.4161***	1.4357***
	(0.4992)	(0.5000)	(0.4998)	(0.4872)	(0.4895)	(0.4899)	(0.4866)	(0.4896)	(0.4887)
Spillovers other	-1.2259***	-1.2020***	-1.2436***	-1.2377***	-1.2268***	-1.2436***	-1.2252***	-1.2141***	-1.2300***
	(0.3805)	(0.3820)	(0.3789)	(0.3748)	(0.3730)	(0.3716)	(0.3731)	(0.3725)	(0.3697)
Fixed effects	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	50070	50070	50070	50070	50070	50070	50070	50070	50070
Firms	3338	3338	3338	3338	3338	3338	3338	3338	3338

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (4)-(6) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. In columns (7)-(9), they are interacted with the average shares over the sample period instead. In columns (4)-(9), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(3), there is no such interactions. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

machinery innovations differently.

Columns (1), (2) and (3) of Table 1.8 reproduce the columns (7), (8) and (9) of Table 1.7 but adding country-year fixed effects, where the country of a firm is still defined as the country with the largest weight. We still obtain a positive effect of low-skill wages on automation innovations with similar elasticities (between 1.8 and 3.0). Columns (4) to (9) go further and only consider the foreign component of wages (and of the other macroeconomic variables). In columns (4) to (9), the foreign low-skill wage variable is defined as the log of the weighted average of country-level wages excluding the domestic country multiplied by the share of foreign low-skill wages in total wages. This share is computed at the beginning of the sample for columns (4) to (6) and as the average value over the whole sample for columns (7) to (9). We pre-multiply the (log) foreign wage by this share to take into allow for some firms being more affected by foreign wages than others, and to ensure that the reported coefficient corresponds to an elasticity on total low-skill wages. The foreign macroeconomic control variables are defined similarly.[34] Once again we

[34]Denote $\omega_{i,D}$ the domestic weight and $\omega_{i,F} = 1 - \omega_{i,D}$ the total foreign weight with $w_{L,D,t}$ the

find a positive effect of low-skill wages on automation innovations, with if anything slightly larger elasticities. Our coefficient captures the average effect of an increase in foreign low-skill wages given our controls whatever the shock behind it. Relative to Table 1.7, the main difference is that high-skill wages are now the macroeconomic control variable with the most explanatory power (neither labor productivity nor GDP per capita have a significant effect once high-skill wages are introduced). Clustering at the country-level (to account for correlation of errors across firms within a country over time) tends to reduce standard errors (Appendix Table 1.8.19). We also reproduced the regression for the auto90 measure, we obtain similar results with slightly smaller coefficients (Appendix Table 1.8.20). Finally, we also replaced the country-year fixed effects with the interaction of country-year dummies with the domestic weight of each firm to account that firms are more or less exposed to the domestic country. Here as well, we obtain similar results although the magnitude of the coefficient on low-skill wages is a bit smaller (Appendix Table 1.8.21).

1.5.3 Non-automation innovations

Is the effect of wages on automation innovations specific to automation or does it affect machinery patents in general? To answer this question, we now look at "placebo regressions" of the effect of wages on innovations with a low score on our automation metric. Specifically, we consider the set of machinery patents and exclude any patent which has a technological category with an automation score above a certain threshold. We fix that threshold at the 60^{th} percentile of the distribution of C/IPC 6 digit codes in the machinery technological fields (0.2091). We refer to these innovations as "placebo machinery" innovations and we recompute knowledge

wage in the domestic country and $w_{L,F,t}$ the average wage in the foreign country. Then we can decompose a small change in $\log w_{L,i,t}$ as:

$$d\log w_{L,i,t} = d\log\left(\omega_{i,D}w_{L,D,t} + \omega_{i,F}w_{L,F,t}\right) = \frac{\omega_{i,D}w_{L,D,0}}{w_{L,i,0}}d\log w_{L,D,t} + \frac{\omega_{i,F}w_{L,F,0}}{w_{L,i,0}}d\log w_{L,F,t}$$

where $\omega_{i,D}w_{L,D,0}/w_{L,i,0}$ denotes the values around which the change is computed—which we take as the the value at the beginning of the period or the average value over the sample period. This shows that if $\frac{\omega_{i,F}w_{L,F,0}}{w_{L,i,0}}d\log w_{L,F,t}$ increases by 0.01 then $w_{L,i,t}$ increases by 1%. The same reasoning applies to high-skill wages or GDP per capita. In (1.3), GDP gap enters directly in levels as an average of logs so we directly interact the domestic and foreign variables with $\omega_{i,D}$ and $\omega_{i,F}$.

stocks and spillover variables for those innovations ("own") and for all innovations except those ("other"). Table 1.9 reports the results. Columns (1) to (3) correspond to the baseline regressions with firm and year fixed effects. Low-skill wages only have a positive and significant effect in column (3) when GDP per capita is included as a control variable, but even in that case the coefficient is statistically significantly smaller than with automation (1.66 versus 3.74 in column 9 of Table 1.7).[35] Columns (4) to (6) repeat the same regressions but add country-year fixed effects and columns (7) to (9) focus on foreign wages (here defined as in columns (4) to (6) of Table 1.8). Neither low-skill wages nor any other macroeconomic control variables have an effect on placebo machinery innovations. The sign of low-skill wages even flip in columns (7) to (9).[36] We therefore view this exercise as validating both our empirical approach and our measure of automation. In particular, if our result on the effect of low-skill wages on automation innovations came from a bias, than that bias would have to be absent for other types of machinery innovations.

1.5.4 Additional results

Innovation types. Building on the previous results contrasting automation innovations and low-automation machinery innovations, we now look at subcategories of automation innovations and a laxer measure in Table 1.10, which reproduces column (8) of Table 1.7 for various types of innovations. Column (1) is essentially a robustness check which removes the codes that we added to the definition of the machinery technological field listed in footnote 11 (though, we continue to exclude the weapons categories). The results are similar to the baseline (with a lower but not statistically so coefficient). Column (2) presents a laxer definition of automation using the 80^{th} percentile of the distribution of the C/IPC 6 digit codes. We still get a positive effect of low-skill wages though with a coefficient smaller than for either auto90 or auto95. Columns (3) to (8) look at subcategories of automation

[35] Further, this positive coefficient in the placebo regression is sensitive to specifications, and unlike for the regressions with automation, it loses significance with different deflators for wages (not shown).

[36] Conditioning on the 60^{th} percentile is not important and we obtain similar results with machinery innovations excluding auto95 or auto90. In fact, since automation innovations are a relatively small share of all machinery innovations, a regression on all machinery innovation gives similar results. See Appendix Table 1.8.22.

Table 1.9: Non-automation innovations

Dependent Variable	Placebo Machinery								
	Domestic + Foreign						Foreign		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	0.2962	0.5837	1.6587**	-0.0486	0.0964	0.6381	-0.7470	-1.0568	-0.9430
	(0.6209)	(0.7013)	(0.6573)	(0.8089)	(0.9245)	(0.9903)	(1.2590)	(1.4477)	(1.3045)
High-skill wage	-0.1907	0.3251	0.8911	-0.3499	-0.0648	0.0238	0.4969	0.1238	0.4016
	(0.6953)	(0.6428)	(0.7506)	(0.9539)	(0.9122)	(1.0053)	(1.3193)	(1.3073)	(1.4470)
GDP gap	-0.0307***	-0.0292***	-0.0292***	-0.0072	-0.0071	-0.0062	0.0117	0.0120	0.0114
	(0.0105)	(0.0103)	(0.0104)	(0.0188)	(0.0187)	(0.0188)	(0.0319)	(0.0319)	(0.0319)
Labor productivity		-1.1140			-0.6087			0.6174	
		(0.7467)			(1.1021)			(1.1452)	
GDP per capita			-3.4367***			-1.5038			0.3079
			(0.8242)			(1.3776)			(1.3051)
Stock own	0.0866**	0.0879**	0.0892**	0.0952**	0.0956**	0.0957**	0.0958**	0.0954**	0.0956**
	(0.0408)	(0.0411)	(0.0405)	(0.0405)	(0.0406)	(0.0404)	(0.0405)	(0.0406)	(0.0406)
Stock other	0.4797***	0.4811***	0.4758***	0.4854***	0.4861***	0.4847***	0.4862***	0.4871***	0.4866***
	(0.0464)	(0.0464)	(0.0463)	(0.0460)	(0.0459)	(0.0459)	(0.0448)	(0.0449)	(0.0449)
Spillovers own	2.6849***	2.7419***	1.9983***	1.1394***	1.1505***	1.0777**	1.1398***	1.1215**	1.1469***
	(0.4153)	(0.4163)	(0.4423)	(0.4410)	(0.4435)	(0.4411)	(0.4393)	(0.4428)	(0.4418)
Spillovers other	-2.4198***	-2.4342***	-1.8132***	-1.2443**	-1.2469**	-1.1918**	-1.2694**	-1.2450**	-1.2706**
	(0.5298)	(0.5348)	(0.5386)	(0.5052)	(0.5056)	(0.5047)	(0.4965)	(0.5008)	(0.4965)
Fixed effects	F + Y	F + Y	F + Y	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	115575	115575	115575	115515	115515	115515	115515	115515	115515
Firms	7705	7705	7705	7701	7701	7701	7701	7701	7701

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Columns (1)-(3) include firm and year fixed effects, while (4)-(9) include firm and country-year fixed effects. Stock variables are calculated with respect to the dependent variable. In columns (7)-(9) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. Domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(6), there is no such interactions. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

innovations. Robot90 and Robot80 were already defined in Section 1.2.5. The other types of innovations are similarly defined: for instance, automat*90 covers patents which belong to technological categories with a frequency of the "automat*" keywords above the threshold used to define auto90. Columns (3) and (4) show that the results are similar for automat* patents (note that by definition automat*80 patents are all auto80 but 91.5% of them are auto90). Column (6) shows that our results extend to robot80 patents (which are also all auto95) but not to robot90 maybe because the sample size is reduced. The sample size drops even more substantially for the CNC categories in columns (7) and (8), and consequently the coefficient on low-skill wages is very imprecisely estimated.

Timing. We look at alternative lags for the macroeconomic and the spillover variables in Table 1.11—we keep a lag of 2 between patent applications and the stocks of patents from the firm because otherwise the dependent variable would be included in the stock of automation when we consider contemporaneous regressions or leads. Column (4) reproduces our baseline results with a 2 year lag. Panel A

Table 1.10: Innovation categories

Dependent Variable	AutoX95	Auto80	Automat* 90	Automat* 80	Robot 90	Robot 80	CNC 90	CNC 80
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low-skill wage	1.9759**	1.3013**	2.6151**	1.7535*	0.4046	2.3998*	-2.6476	-1.5273
	(0.9046)	(0.6373)	(1.1768)	(0.9657)	(1.6931)	(1.2440)	(2.0151)	(1.5877)
High-skill wage	-1.2113	-1.2776**	-0.9885	-0.9874	-0.8384	-2.0705*	2.0374	0.8833
	(0.9265)	(0.5754)	(1.0579)	(0.8395)	(1.6053)	(1.2334)	(1.8923)	(1.5580)
GDP gap	0.0370**	0.0047	0.0078	-0.0052	0.0345	0.0409	0.0317	0.0214
	(0.0186)	(0.0121)	(0.0214)	(0.0173)	(0.0365)	(0.0264)	(0.0411)	(0.0305)
Labor productivity	0.2216	0.8058	-0.9351	-0.2196	0.8059	0.7937	2.7221	1.9101
	(0.9431)	(0.6648)	(1.1098)	(0.9161)	(1.9404)	(1.3971)	(2.3494)	(2.1381)
Stock own	-0.1400**	0.0263	-0.1149*	-0.0861	-0.3029***	-0.1319*	-0.3043**	-0.2888***
	(0.0567)	(0.0374)	(0.0601)	(0.0525)	(0.0993)	(0.0790)	(0.1511)	(0.0999)
Stock other	0.6443***	0.5225***	0.6684***	0.6312***	0.8200***	0.6329***	0.5642***	0.6140***
	(0.0645)	(0.0460)	(0.0872)	(0.0737)	(0.1334)	(0.0994)	(0.1303)	(0.0961)
Spillovers own	0.7068*	0.9236*	0.3869	0.4415	0.2346	0.1891	0.7408**	0.4634*
	(0.4072)	(0.5235)	(0.4365)	(0.4719)	(0.5380)	(0.3489)	(0.3657)	(0.2727)
Spillovers other	-0.5863*	-0.6139	-0.3800	-0.3305	-0.0665	-0.2028	-1.5340***	-0.7109
	(0.3036)	(0.4435)	(0.2736)	(0.3469)	(0.3529)	(0.2887)	(0.5522)	(0.4478)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	48600	97635	34170	50220	17670	24645	8970	15000
Firms	3240	6509	2278	3348	1178	1643	598	1000

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Stocks and spillovers are calculated with respect to the dependent variable. All regressions include firm fixed effects and year dummies. All regressions include a dummy for no stock and no spillover. Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

shows that the largest coefficient on low-skill wages is obtained for a 2 year lag, but remains relatively stable between a 4 year lag and a 1 year lead. Both panels find an effect of low-skill wages more clearly centered around lag 2 (ADHMV also found that the largest coefficient for the effect of gas prices on innovations in the car industry was at a 2 year lag). Our baseline regressions assume a 2 year lag between wages and patent applications.

Of course, innovators would not be interested about wages 2 years in the past per se, but only inasmuch as they are indicative of future wages. This is our interpretation throughout our regressions, with the 2 year lag corresponding roughly to the time spent between an effect on R&D and the first results materialized by a patent application.[37] We push this logic further in Appendix Table 1.8.23, where we compute predicted future wages at time $t-2$ based on an AR(1) process with country-specific trends. We find similar results.

Minimum wage. Given its policy relevance, we also look at the effect of minimum wages using data on 22 countries.[38] Importantly, we cannot use the minimum wage

[37] In that context, the difference between the significant lead coefficients in Panel A and the insignificant ones in Panel B and C, could reflect that domestic wages may be easier to predict than foreign wages.

[38] We use data from the OECD. Importantly, not all countries have government-mandated minimum wages, most notably Italy and, until 2015, Germany. For Germany, we follow Dolado,

Table 1.11: Lags and leads

Dependent variable	Auto95							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lags (Leads)	-5	-4	-3	-2	-1	0	1	2
Panel A: baseline								
Low-skill wage	1.4268*	2.0578**	1.9681**	2.4266***	2.0882**	2.0767**	2.2411***	1.4514*
	(0.8599)	(0.8328)	(0.8229)	(0.8658)	(0.8417)	(0.8331)	(0.8518)	(0.8251)
High-skill wage	-0.0640	-0.9379	-1.6808*	-1.6700*	-2.0273**	-2.5752***	-2.5365***	-2.7223***
	(0.9033)	(0.8937)	(0.9223)	(0.8634)	(0.7977)	(0.8281)	(0.7687)	(0.7828)
Labor productivity	0.1931	0.4055	1.1283	0.1285	0.0857	-0.0118	-0.2255	0.4201
	(1.1023)	(1.0789)	(1.0884)	(0.9199)	(0.7871)	(0.8022)	(0.8265)	(0.8912)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	47565	48240	49395	50115	50670	51315	52470	53940
Firms	3171	3216	3293	3341	3378	3421	3498	3596
Panel B: country-year fixed effects								
Low-skill wage	0.9671	1.3572	1.5405	2.1429*	1.6930	1.2360	1.2538	0.1282
	(1.1012)	(1.1353)	(1.1175)	(1.1505)	(1.1222)	(1.1088)	(1.1409)	(1.0962)
High-skill wage	0.4539	-0.9749	-1.7245	-1.9117*	-2.0866**	-2.7165**	-2.1045**	-1.6862
	(1.3522)	(1.1490)	(1.0931)	(1.0157)	(1.0346)	(1.0935)	(1.0333)	(1.0682)
Labor productivity	-1.5193	-0.8311	-0.2556	-1.2851	-0.5775	0.3167	-0.1957	0.0676
	(1.8190)	(1.6338)	(1.5444)	(1.6381)	(1.6431)	(1.5761)	(1.6158)	(1.5974)
Panel C: country-year fixed effects and foreign variables								
Low-skill wage	1.5679	2.5117*	3.1804**	4.3023***	3.0459**	1.6943	1.6996	0.4034
	(1.6579)	(1.4908)	(1.4684)	(1.4482)	(1.4516)	(1.5642)	(1.7055)	(1.7377)
High-skill wage	2.1192	-1.0194	-2.5135	-2.4740*	-3.2862**	-3.8818***	-3.3215**	-2.5666*
	(1.8327)	(1.6302)	(1.6445)	(1.4209)	(1.4238)	(1.4272)	(1.3771)	(1.4844)
Labor productivity	-2.3858	-0.9029	-0.7200	-1.7494	0.4010	1.8684	1.6417	1.6644
	(1.5235)	(1.5420)	(1.5937)	(1.4131)	(1.3247)	(1.4493)	(1.5255)	(1.6175)
Fixed effects	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	47565	48240	49365	50070	50595	51255	52410	53895
Firms	3171	3216	3291	3338	3373	3417	3494	3593

Note: Marginal effects; Standard errors in parentheses. Each panel represents a different regression. All regressions contain controls for GDP gap, stocks and spillovers, for which we do not report the coefficient. The independent variables (wages, VAemp and GDP gap) are lagged by the number of periods indicated in lag, except for the stock variables which are always lagged by 2 periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Panel A regressions contain firm and year fixed effects. Panel B and C regressions contain firm and country-year fixed effects. In Panel C regressions, wages are replaced with foreign wages interacted with the share of foreign wages in total wages at the beginning of the sample, and similarly for the other macro variables. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

as an instrument for low-skill wages: our regressions show that the high-skill wage has a significant effect and therefore should be included in the regression. If low-skill wages were to be instrumented so should high-skill wages, and we would need a second instrument. We report the results of reduced form regressions where we replace low-skill wages with the minimum wage in Appendix Table 1.8.24. We find a positive effect of the minimum wage on automation innovations which in specifications with country-year fixed effects has p-values just above or below 0.1. Clustering standard-errors at the country-level gives significant coefficients (see Appendix Table 1.8.25). Minimum wages are unlikely to be a strong predictor of automation in our analysis: first because it only captures part of the labor costs (contrary to our baseline measure), second because we focus on automation innovations that largely happen in manufacturing where wages for low-skill workers are often substantially higher than the minimum wage. An analysis on automation in service industries might show a stronger relationship.

1.5.5 Shift-share set-up

A recent literature addresses the identifying assumptions behind the shift-share set-up in linear regressions. In the language of our setting, Goldsmith-Pinkham, Sorkin and Swift (2019) show that the shift-share instrument is equivalent to a combination of weights time country-year dummies. Our shift-share setting would then capture the effect of low-skill wages on automation innovations if weights time country-year dummies only affect automation through the controls that we have included. In this interpretation of the shift-share set-up the exogeneity of the weights is important and we show below that our results are robust to using weights from an earlier period.

Borusyak, Hull and Jaravel (2018) show that country-time shocks can also be a source of identification in the shift-share setting. The inference is valid if either there is a large number of countries (such that the Herfindahl index tends toward 0) affected by independent shocks (controlling for a year and firm fixed effects); or the correlation of shocks within a country decays sufficiently rapidly that a large

Kramarz, Machin, Manning, Margolis, Teulings, Saint-Paul and Keen (1996) and use the the collectively bargained minimum wages which in effect constitute law.

number of country x years is sufficient (see Appendix A2 in their paper). They advise practitioners to use appropriate controls to capture omitted variables. We follow this approach partly by including a large set of controls in our regressions and partly by including country-year fixed effects. They further encourage practitioners to apply the standard error correction of Adão, Kolesár and Morales (2019).

Adão et. al. (2019) show that standard applications with the shift-share design often lead to an over-rejection of the null of no effect. In the language of our application, the problem arises if the standard errors of firms with similar country-distributions have correlated residual errors. Though this problem is related to the correlation of standard errors in clustered designs it is not solved by standard clustering. Adão et. al. (2019) use Monte Carlo simulations in a standard Bartik setting and show that in a setting where the true coefficient is zero by construction the commonly used approach rejects the null of no effect up to 55% of the cases. They derive a formulae for standard errors in an OLS setting that corrects for this problem. This formulae is not directly applicable in the current setting since we employ a Poisson estimator and deriving the corresponding correction for the Poisson estimator is beyond the scope of this paper. Instead we implement a Monte Carlo simulation like the one used in Adão et. al. (2019) and show that we do not have a similar problem of over-rejection.

Specifically, we replicate the regression in Column (9) in Table 1.7. For each firm we keep the automation activity, the stocks of innovations, the spillover variables, as well as the distribution of country-weights based on actual patents. For each country we sample without replacement the entire path of wages and GDP from the existing set of wages and GDP. Figure 1.5 shows a histogram of the t-statistic of the low-skill wages, the main coefficient of interest, where the red line at 4.1 corresponds to the t-statistic of column (9). The empirical distribution has a heavier left tail than the expected standard normal distribution but only in 0.33% of the cases did the absolute value of the t-statistic exceed the realized t-statistic from our main regressions. In the language of Adão et al. (2019) the set of controls soaks up most country-specific shocks affecting the outcome variable and, consequently, no shift-share structure is left in the regression residuals.

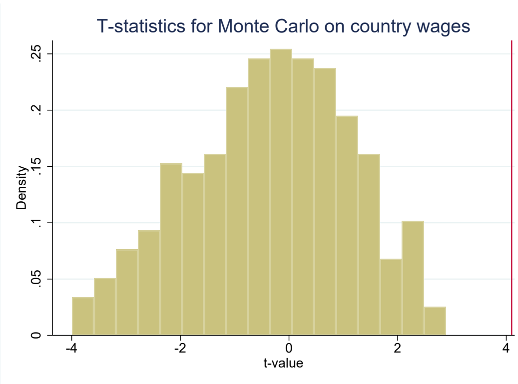


Figure 1.5: The t-value on the low-skill wages from a Monte Carlo simulation sampling country wages and GDP (see text for details)

1.5.6 Robustness checks

This section presents several robustness checks.

Controlling for the cost of innovation. Our measure of wages could still reflect the cost of innovation if innovation does not solely take place in the domestic country. To address this issue we re-build our firm-specific macroeconomic variable using the inventor weights of the firm instead of the patent weights. Table 1.12 reports the result. The coefficient on low-skill wages remains positive and significant but the coefficient on low-skill wages weighted by inventor weights is small and insignificant. These regressions constitute a placebo test in that they are essentially treating the firms by the same macroeconomic shocks but distributed according to where the firm innovates not where it sells.

Multicollinearity and skill premium. Low-skill wages, high-skill wages and labor productivity are correlated, which could affect our regressions—although controlling for firm fixed effects and year fixed effects, the correlation coefficient is only 0.6 (see Appendix Table 1.8.26). To deal with this issue, Appendix Table 1.8.27 regresses automation innovation on the log of the ratio of low-skill to high-skill wages (the inverse of the skill premium) for firm fixed effects, country-year fixed effects and foreign wages with country-year fixed effects. The coefficient on the inverse skill premium is always of the same magnitude as that on low-skill wages and highly significant. On the other hand, replacing low-skill and high-skill wages with their ratio in the regressions with placebo machinery innovations of Table 1.9

Table 1.12: Wages weighted by inventor weights

Dependent Variable	Auto95								
	Domestic + Foreign						Foreign		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	2.6194***	2.4897***	3.7088***	1.9136*	2.0761*	2.9547**	4.7342***	5.6526***	5.0494**
	(0.9119)	(0.9549)	(1.0503)	(1.0705)	(1.1954)	(1.3173)	(1.5977)	(1.7376)	(1.9638)
Low-skill wage (iw)	-0.2924	-0.1985	-0.0762	-0.1439	0.0552	0.0944	-0.1005	0.6363	0.4886
	(0.4461)	(0.4668)	(0.4805)	(0.4747)	(0.4794)	(0.4754)	(0.5772)	(0.6011)	(0.5562)
High-skill wage	-1.9307**	-2.1087**	-0.8557	-2.5728**	-2.2029**	-1.9204*	-4.0721***	-3.4857**	-4.2454**
	(0.9171)	(1.0032)	(0.8490)	(1.0770)	(1.0546)	(1.1427)	(1.5497)	(1.6359)	(1.6521)
High-skill wage (iw)	0.3960	0.5295	0.4991	0.1804	0.4874	0.2735	-0.2895	0.7720*	0.0817
	(0.3397)	(0.3869)	(0.3370)	(0.3249)	(0.3727)	(0.3451)	(0.4384)	(0.4655)	(0.4573)
GDP gap	0.0364	0.0366	0.0314	0.0616*	0.0616*	0.0630*	-0.0077	-0.0166	-0.0080
	(0.0229)	(0.0227)	(0.0231)	(0.0362)	(0.0362)	(0.0361)	(0.0567)	(0.0565)	(0.0565)
GDP gap (iw)	-0.0076	-0.0083	-0.0050	0.0003	-0.0017	0.0022	0.0186	0.0126	0.0208
	(0.0123)	(0.0121)	(0.0124)	(0.0122)	(0.0120)	(0.0125)	(0.0152)	(0.0153)	(0.0153)
Labor productivity		0.4313			-0.8383			-1.5747	
		(1.1116)			(1.6547)			(1.5093)	
Labor productivity (iw)		-0.3065			-0.7076			-1.8365***	
		(0.5374)			(0.5066)			(0.6146)	
GDP per capita			-3.0004***			-2.4889			-0.1854
			(0.9236)			(1.9888)			(2.1553)
GDP per capita (iw)			-0.4388			-0.4514			-1.1406**
			(0.5746)			(0.6508)			(0.5649)
Control variables	stock + spill	stock + spill	stock + spill	stock + spill	stock + spill	stock + spill	stock + spill	stock + spill	stock + spill
Fixed Effects	F + Y	F + Y	F + Y	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	49305	49305	49305	49245	49245	49245	37395	37395	37395
Firms	3287	3287	3287	3283	3283	3283	2493	2493	2493

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Columns (1)-(3) include firm and year fixed effects, while (4)-(9) include firm and country-year fixed effects. Stock variables are calculated with respect to the dependent variable. In columns (7)-(9) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and labor productivity. Domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(6), there is no such interactions. All regressions with patent-weighted low-skill wage variable include a corresponding inventor-weighted low-skill wage variable, similarly for high-skill wage, GDP gap, GDP per capita and labor productivity. All inventor-weighted variables are denoted by (iw) after their names. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

gives insignificant coefficients.

Including middle-skill wages. Previous work has often found that IT disproportionately affects middle-skill workers (e.g. Autor and Dorn, 2013). In this spirit, Appendix Table 1.8.28 adds middle-skill wages to our regressions. Low-skill wages continue to have a large positive impact on automation, whereas middle-skill wages have a negative (though not consistently significant effect). Low-skill and middle-skill wages are highly correlated (with a coefficient of 0.94, see Appendix Table 1.8.26), and consequently, middle-skill wages have a positive coefficient when low-skill wages are not included.

Wages and deflators. Appendix Table 1.8.29 shows that our results are robust to deflating our macroeconomic variables differently: by converting to USD in a different year (columns 1 and 2), every year (columns 3 and 4) or using the local GDP deflator instead of the local PPI in manufacturing (columns 5 and 6). Further, we look at total wages instead of manufacturing wages either with our baseline deflator (columns 7 and 8) or converting every year (columns 9 and 10). Our results remain largely robust but with smaller coefficients when converting in USD every year. Converting in USD every year makes our macroeconomic variables more correlated and increases the importance of short-term fluctuations.

Weights. Table 1.13 reproduces the regressions of columns (7) and (8) of Table 1.7 but with alternative firm-specific weights $\omega_{i,c}$. In columns (1) and (2), we compute the patent weights over a more recent period (1985-1994 instead of 1970-1994) and obtain the same results. Columns (3) and (4) on the other hand drop the 5 most recent years in computing the weights. We lose a large number of firms, but still obtain a positive effect of low-skill wages on automation innovations, though with slightly smaller coefficients. This regression addresses the potential concern that our weights could be endogenous because firms which already intend to do automation innovations may decide to locate in places where they forecast an increase in low-skill wages: it is hard to see how firms' location decisions before 1989 could reflect increases in wages from 1995 onward.[39] Columns (5) to (10) keep the patent weights as in our baseline analysis but instead of multiplying them by $GDP_c^{0.35}$, they do not multiply them (columns 5 and 6), multiply them by GDP (7 and 8) or by the total value of low-skill employment to the power 0.35 ($(w_L L)^{0.35}$: one could argue that this represents a better measure than $GDP^{0.35}$ of the potential market for technology designed to automate low-skill work). We obtain similar results.

Quality. Appendix Table 1.8.30 investigates whether our results are robust when focusing on patents of higher quality. We look at patents which have been applied for at 2 of the 3 main patent offices (EU, Japan and US), or at triadic patents which have been applied for at the 3 offices. Triadic patents are generally considered to be patents of very high quality. All of these give similar results. We also restrict attention to biadic patents with at least one citation within 5 years and weigh patents by citations.[40] This weakens the results somewhat perhaps because whereas the decision to innovate is a choice variable of the firm the eventual quality of the innovation is largely random.

[39]The same concern can be addressed by keeping our baseline weight but dropping the first few years. See Appendix Table 1.8.30 which reproduces Table 1.7 but only from 2000. Though the standard errors are bigger, the results are essentially the same.

[40]We add to each patent the number of citations received within 5 years normalized by technological field and year of application, in a similar fashion to, for instance, Kogan, Papanikolaou, Seru and Stoffman (2017), who find a positive correlation between patent value and citations. Abrams, Akcigit and Grennan (2018) on the other hand find an inverted U relationship between patent value and citations.

Table 1.13: Alternative weights

Dependent Variable	Auto95									
	1985 – 1994		1970 – 1989		GDP^0		GDP^1		$(w_L * L)^{0.35}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Low-skill wage	2.4739***	2.3626***	1.8155*	1.7192*	1.8685**	1.7962**	2.8690***	2.8825***	2.2007***	2.1429**
	(0.8691)	(0.8876)	(0.9480)	(0.9544)	(0.7776)	(0.8176)	(0.8855)	(0.8953)	(0.8125)	(0.8516)
High-skill wage	-1.7055**	-1.9002**	-0.8990	-1.0259	-1.3791*	-1.4820*	-1.6609**	-1.6405**	-1.4445*	-1.5237*
	(0.8288)	(0.8899)	(0.8354)	(0.9524)	(0.8226)	(0.8851)	(0.7114)	(0.7547)	(0.7847)	(0.8444)
GDP gap	0.0226	0.0224	0.0140	0.0138	0.0276*	0.0273*	0.0265*	0.0264*	0.0283*	0.0280*
	(0.0163)	(0.0162)	(0.0164)	(0.0164)	(0.0154)	(0.0153)	(0.0158)	(0.0159)	(0.0156)	(0.0154)
Labor productivity		0.4484		0.3240		0.2559		-0.0482		0.1983
		(0.9649)		(1.0211)		(0.8994)		(0.8293)		(0.9221)
Stock automation	-0.1337**	-0.1343**	-0.1194**	-0.1201**	-0.1436***	-0.1441***	-0.1429***	-0.1429***	-0.1428***	-0.1432***
	(0.0524)	(0.0524)	(0.0602)	(0.0603)	(0.0509)	(0.0511)	(0.0511)	(0.0511)	(0.0509)	(0.0509)
Stock other	0.6539***	0.6540***	0.6900***	0.6895***	0.6414***	0.6410***	0.6385***	0.6384***	0.6404***	0.6403***
	(0.0639)	(0.0639)	(0.0769)	(0.0769)	(0.0600)	(0.0600)	(0.0598)	(0.0598)	(0.0600)	(0.0600)
Spillovers automation	0.5655*	0.5815*	0.2618	0.2719	0.4091	0.4178	0.8056**	0.8051**	0.4679	0.4744
	(0.3154)	(0.3182)	(0.3206)	(0.3229)	(0.3093)	(0.3106)	(0.3340)	(0.3354)	(0.3103)	(0.3114)
Spillovers other	-0.3401	-0.3693	-0.3772	-0.3951	-0.1913	-0.2090	-0.4680**	-0.4664**	-0.2577	-0.2702
	(0.2303)	(0.2401)	(0.2435)	(0.2518)	(0.2311)	(0.2366)	(0.2265)	(0.2305)	(0.2284)	(0.2353)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	45735	45735	35955	35955	50115	50115	50115	50115	50115	50115
Firms	3049	3049	2397	2397	3341	3341	3341	3341	3341	3341

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). In columns (1) and (2) firms' country weights for the macroeconomic variables are computed over the period 1985-1994; and over the period 1970-1989 for columns (3) and (4). Columns (5) to (19) use the baseline pre-sample period of 1970-1994. Columns (5) and (6) do not adjust for GDP in the computation of the weights and columns (7) and (8) use GDP instead of $GDP^{0.35}$ to adjust for countries' size in the computation of the weights. Columns (9) and (10) adjust for total low-skilled payment instead of using GDP. Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nickell's bias. Our regressions include the stock of automation innovations and therefore may suffer from Nickell's bias. To address this issue, in Appendix Table 1.8.31, we remove the stock of automation innovations with very little effect on the variable of interest. In addition, we use Blundell, Griffith and van Rennen (1999)'s method, which proxies for the fixed effect by using the firm's pre-sample average of the dependent variable. We obtain very similar results.

Industry-year fixed effects. Appendix Table 1.8.33 introduces industry-year fixed effects where the industry of a firm correspond to its 2 digit industry in Orbis. The results are very similar.

1.6 Event study: the Hartz reforms in Germany

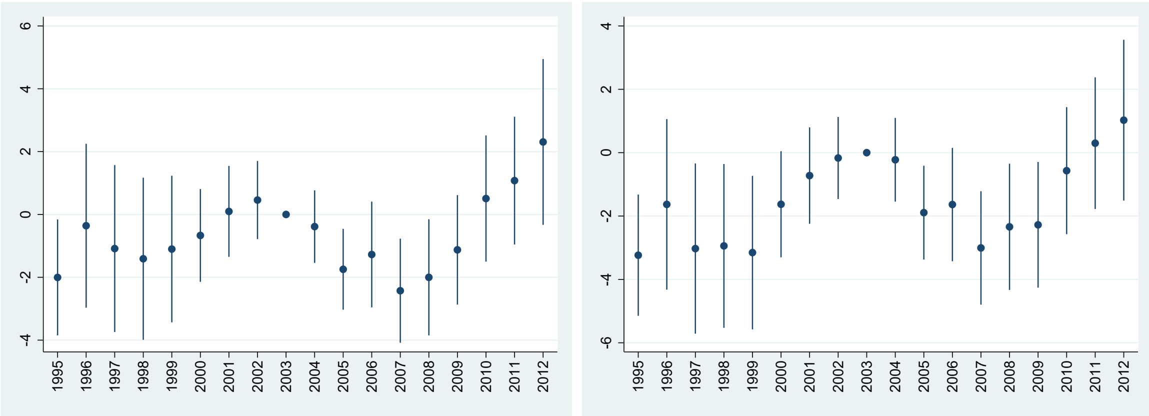
In this section, we use the Hartz reform as an event study to complement our main analysis. The Hartz reforms were a series of labor-market reforms in Germany designed from 2002 onward and implemented between January 1st 2003 and January 1st 2005. The reforms aimed at reducing unemployment and increasing labor-market flexibility by reforming employment agencies to provide better job-search assistance, deregulating temporary work, offering wage subsidies for hard-to-place workers, reducing or removing social contributions on low-paid jobs and reducing long-term unemployment benefits (see Jacobi and Kluve, 2007). The reforms have been widely credited with playing a major role in the remarkable performance of the German labor market since, in particular, for increasing labor supply and improving matching efficiency (see Krause and Uhlig, 2012, Krebs and Scheffel, 2013 and 2017, or Burda and Seele, 2016). Such reforms should reduce the incentive to automate low-skill labor by reducing labor costs (directly through social contribution and indirectly through an increase in labor supply) but also by allowing for more flexible contracts and reducing the expected cost of vacancies.

We start from the same database linking firms and patents as in our main empirical analysis of Section 1.5, using the same weights to measure firms' exposure to different countries and focusing on biadic patent applications as a measure of innovation. We still define the country of a firm as the country of largest weight, and restrict attention to firms from the countries with the highest average exposure to Germany (Austria, France, Italy, Japan, the Netherlands, Spain, Switzerland, the United Kingdom and the United States).

We first run the following Poisson regression, over the years 1995-2012, maintaining a 2-year lag:

$$PAT_{Aut,i,t+2} = \exp\left(\beta_{DE} \cdot \delta_t \omega_{i,DE} + \beta_{Ka} \ln K_{Aut,i,t} + \beta_{Ko} \ln K_{other,i,t} + \delta_i + \delta_{c,t}\right) + \epsilon_{i,t}.$$

$PAT_{Aut,i,t+2}$ is still a count of automation patents, $K_{Aut,i,t}$ and $K_{other,i,t}$ continue represent firm knowledge stocks, δ_i is a firm fixed effect, $\delta_{c,t}$ is a country-year fixed effect, $\omega_{i,DE}$ is the (fixed) firm weight on Germany, δ_t is a set of year dummies (with 2003 as the excluded year) and β_{DE} is the full vector of coefficients of interest. The vector of coefficients β_{DE} determines by how much a firm more exposed to Germany tends to do more automation patents in a given year relative to 2005 (with the 2 year lag). Figure 1.6.a reports the results. The coefficients can be interpreted as follows: a value of -2 in 2008 indicates that on average a firm with a German weight of 0.1 (the mean value is 0.106) did 20% less automation innovations in 2010 than in 2005 (recall the 2 year lag) compared to a firm with no German exposure.



(a) Effect of German exposure on automation innovations

(b) Effect of German exposure on automation innovations relative to other machinery innovations.

Figure 1.6: Effect of German exposure on automation innovations. Panel (a) reports coefficients on the interaction between the German weight and a set of year fixed effects in a Poisson regression of auto95 innovations controlling for a full set of fixed effects and firm innovation stocks. Panel (b) reports coefficients on the triple interaction between the German weight, a dummy for auto95 innovations and a set of year fixed effects in a Poisson regression of auto95 and other machinery innovations controlling for a full set of fixed effects, firm innovation stocks and the interaction between the German weight and a set of year fixed effects.

The figure suggests that from 2000 until 2004 firms highly exposed to Germany increased their propensity to introduce automation innovations. This trend reversed between 2006 and 2009 and resumed from 2010. This is consistent with the Hartz reform increasing labor supply from 2002-2004, and therefore decreasing the incentive to introduce automation innovations 2 years later. 2008 marks the beginning of the Great Recession which had a lower impact on German labor markets than in other countries, so that German labor markets remained relatively tight, potentially increasing the incentive to undertake automation innovations.

The previous figure clearly shows that the behavior of firms highly exposed to Germany differs over time from that of other firms. To show that the trends above are specific to automation innovations, we then run the following Poisson regression:

$$PAT_{k,i,t+2} = \exp\begin{pmatrix} \beta_{DE} \cdot \delta_t \omega_{i,DE} + \beta_{DE}^{aut} \cdot \delta_t \omega_{i,DE} 1_{k=aut} \\ +\beta_{Ka} \cdot \delta_k \ln K_{Aut,i,t} + \beta_{Ko} \cdot \delta_k \ln K_{other,i,t} + \delta_{k,i} + \delta_{k,c,t} \end{pmatrix} + \epsilon_{k,i,t}. \tag{1.5}$$

k denotes the type of an innovation which is either auto95 or another machinery innovation, $\delta_{k,i}$ represents a full set of innovation type firm fixed effects and $\delta_{k,c,t}$ innovation type country year fixed effects and $1_{k=aut}$ is a dummy for an auto95 innovation. Standard errors are clustered at the firm level. β_{DE}^{aut} is the vector of coefficients of interests, for each year, they measure how much exposure to Germany increases the relative propensity to introduce automation innovations instead of other forms of machinery innovations compared to 2005 (as 2003 is still the excluded year). Figure 1.6.b reports the results: the pattern is similar to Figure 1.6.a but more striking.

To formally test that the Hartz reform created a trend break in the relative propensity of firms highly exposed to Germany to introduce automation innovations relative to other types of machinery innovation, we replace the full set of year fixed-effects δ_t in $\beta_{DE}^{aut} \cdot \delta_t \omega_{i,DE} 1_{k=aut}$ in (1.5) with a time trend $t - 2003$ and a time trend interacted with a post 2003 dummy $(t-2003)1_{t>2003}$. We focus on the years 1998-2008 (i.e. 2000-2010 for the innovation variable) to have a panel centered on 2003 and avoid the Great Recession. Table 1.14 reports the result. Column (2) corresponds exactly to the specification we discussed: it shows a significant time trend in the effect of German exposure on the relative propensity to carry automation innovation two years later between 1998 and 2003, this time trend sharply reversed in the following 5 years. Column (1) carries out the same regression but omits the controls for the stock variables. Column (3) adds a control for the triple interaction of the German weight, a dummy for automation innovations and a dummy for post-2003. This tests whether the break in time trends is associated with a shift in levels. The coefficient is insignificant. Column (4) replaces the German weight by a dummy indicating that the firm is in the top quartile of exposure to Germany among innovating firms: the results are of similar magnitude since the 75^{th} percentile of German weight is 0.16. Column (5) uses the low-automation innovations of section 1.5.3 instead of all other machinery innovations. The results are similar. Finally, column (6) considers three types of innovations by separating non-auto95 machinery innovations into the low-automation innovations of the previous columns and the rest. There are no significant trends distinguishing low-automation innovations from other non-auto95 machinery innovations. Overall, this exercise suggests

Table 1.14: Innovation and exposure to Germany

Dependent variables	Auto 95 and other + low auto				Auto95 and low auto	Auto95 and other and low auto
	(1)	(2)	(3)	(4)	(5)	(6)
time trend*dummy auto95*German exposure	0.6309**	0.6245***	0.7726*	0.0929**	0.6486***	0.6523***
	(0.2502)	(0.2296)	(0.3957)	(0.0366)	(0.2464)	(0.2322)
time trend*dummy auto95*post_2003*German exposure	-1.2330***	-1.2322***	-1.3229**	-0.1810**	-1.2500***	-1.2826***
	(0.4473)	(0.4291)	(0.5273)	(0.0766)	(0.4605)	(0.4300)
dummy auto95*post_2003*German exposure			-0.7289			
			(1.0856)			
time trend*dummy low auto*German exposure						0.0081
						(0.1278)
time trend*dummy low auto*post_2003*German exposure						-0.0386
						(0.1835)
year dummy*German exposure	Y	Y	Y	Y	Y	Y
firm innovation stocks * innovation types	N	Y	Y	Y	Y	Y
firm *innovation types fixed effects	Y	Y	Y	Y	Y	Y
country * year * innovation types fixed effects	Y	Y	Y	Y	Y	Y

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm innovation types fixed effects, country year innovation types fixed effects and controls for the year dummy times the measure of German exposure. German exposure is measured by the German weights in all regressions except for column (4) where it is replaced by a dummy signaling that the firm is in the top quartile of Germany exposed firms. Innovation types are auto95 and (other + low auto) in columns (1) to (4), auto 95 and low auto in column (5) and auto 95, other and low auto in column (6). All regressions with stock variables include a dummy for no stock. Standard errors are clustered at the firm-level.* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

that the Hartz reforms reduced the propensity of foreign firms highly exposed to Germany to introduce automation innovations.

1.7 Conclusion

In this paper, we have used patent text data to identify patents which correspond to automation innovations and provide a new measure of automation. Across sectors, our measure is uncorrelated with computerization but positively correlated with robotization. We also find that our measure is associated with a decline in routine tasks across US sectors. We then use our classification to analyze for the first time the effect of wages on automation innovations in machinery. We find that automation innovations are very responsive to changes in low-skill wages with elasticities estimated between 2 and 4. This result does not extend to other types of innovations in machinery. Furthermore, we show that the Hartz reforms in Germany were associated with a relative increase in automation innovations by foreign firms with a high exposure to Germany.

These results suggest that policies which increase labor costs for low-skill workers will lead to an increase in innovations which aim at saving on low-skill workers. Therefore, with endogenous technological change, such policies are likely to be less costly for the economy in terms of overall welfare, but they introduce additional negative effects for low-skill workers. By how much then would an exogenous increase

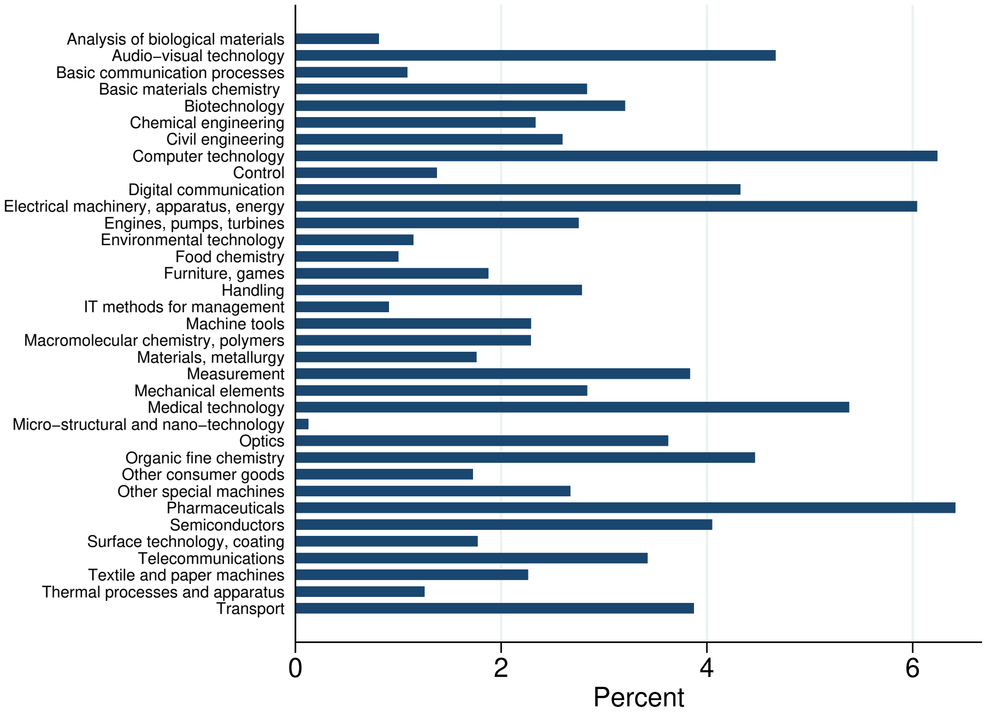
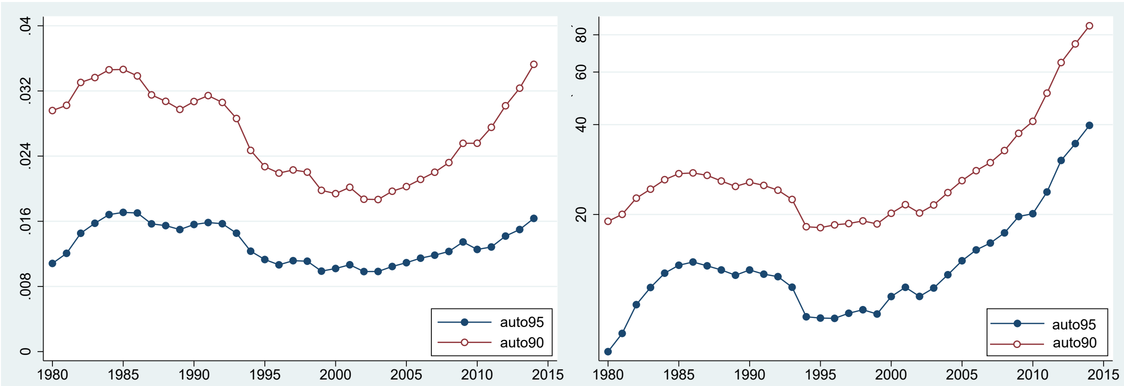


Figure 1.8.7: Share of biadic patent applications in the different technical fields in 1997-2011

in low-skill wages be undone in a couple of years through innovation? Answering this question requires finding the effect of an increase in automation patents on low-skill wages.

1.8 Appendix

1.8.1 Further Figures and Tables



(a) Share of automation patents in machinery out of total patents. Automation technological categories are defined at the 90th percentile of the distribution of 6 digit C/IPC codes in machinery (for auto90) or the 95th percentile (auto95).

(b) Number of automation patents worldwide according to the auto90 and auto95 definitions

Figure 1.8.8: Trends in automation (for biadic applications)

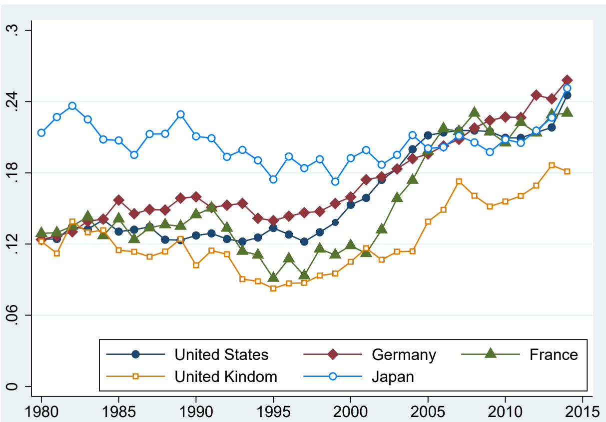


Figure 1.8.9: Share of automation patents (auto95) in machinery by applicant's nationality.

Table 1.8.15: Share of automation patents in machinery across sectors

ISIC Rev. 4	Title	Share of automation patents in machinery 1997 - 2011 (in %)					
		Germany		United States		All Countries	
		auto95	auto90	auto95	auto90	auto95	auto90
A	Agriculture, forestry and fishing	5.7	12.4	6.4	14.8	6.8	13.8
B	Mining and quarrying	10.0	17.6	9.9	18.2	9.8	17.2
10-12	Food, beverages and tobacco products	4.6	12.9	5.6	15.2	5.0	12.6
13-15	Textiles, wearing apparel, leather and related products	3.9	9.0	4.7	11.4	4.2	10.3
16	Wood and products of wood and cork	4.3	9.3	4.7	11.9	4.9	10.9
17-18	Paper, paper products and printing	2.6	6.8	2.8	7.5	2.8	7.6
19-22	Coke, chemicals, pharmaceuticals, rubber and plastic products	2.9	6.9	3.8	8.2	3.0	7.0
23	Other non-metallic mineral products	6.1	11.7	6.7	13.9	5.9	12.0
24	Basic metals	10.8	26.0	12.4	29.4	11.1	27.0
25	Fabricated metal products	7.7	22.3	8.8	24.3	8.4	23.7
26-27	Computer, electronic, optical and electrical products	30.7	39.4	30.1	40.1	29.4	39.1
28	Machinery and equipment n.e.c.	17.4	30.5	18.1	30.7	18.8	31.5
29	Motor vehicles, trailers and semi-trailers	32.6	36.8	30.0	35.7	31.9	36.8
30	Other transport equipment	24.5	29.3	22.8	29.1	26.1	31.9
31	All other manufacturing branches	15.7	23.2	18.7	27.9	18.9	27.7
E	Water supply; sewerage, waste management and remediation activities	6.6	13.2	8.2	16.5	7.9	14.7
F	Construction	7.7	11.7	9.4	15.5	8.4	13.3

Table 1.8.16: Top 10 auto95 innovators in our sample

Company	Number of biadic auto95 patents in 1997-2011
Siemens S.A.	1738
Honda Motor Co., Ltd.	810
Fanuc Co.	777
Samsung Electronics Co., Ltd.	706
Robert Bosch AG	655
Mitsubishi Electric Europe B.V.	652
Tokyo Electron Europe, Ltd.	578
Murata Machinery, Ltd.	501
Kabushiki Kaisha Toshiba	473
General Electric Canada	464

Table 1.8.17: Baseline regressions for auto95 with country-level clustering

Dependent variable	Auto95								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	2.2000***	2.8254***	1.8160***	1.9058***	1.9992**	2.2954***	2.4627***	2.4266***	3.7365***
	(0.5464)	(0.7421)	(0.6310)	(0.6863)	(0.9001)	(0.5383)	(0.7170)	(0.8727)	(0.6582)
High-skill wage		-0.9210	-0.9009**	-0.9695***	-0.8698	-0.2971	-1.6180***	-1.6700**	-0.4838*
		(0.6234)	(0.3519)	(0.3701)	(0.7025)	(0.2972)	(0.4701)	(0.7968)	(0.2831)
Stock automation			-0.1275***	-0.1269***	-0.1270***	-0.1239***	-0.1441***	-0.1443***	-0.1504***
			(0.0336)	(0.0339)	(0.0335)	(0.0355)	(0.0358)	(0.0365)	(0.0389)
Stock other			0.6311***	0.6296***	0.6309***	0.6260***	0.6408***	0.6407***	0.6489***
			(0.0495)	(0.0506)	(0.0483)	(0.0518)	(0.0493)	(0.0492)	(0.0501)
GDP gap				0.0210***	0.0214**	0.0179**	0.0279***	0.0278***	0.0265***
				(0.0081)	(0.0088)	(0.0074)	(0.0091)	(0.0096)	(0.0076)
Labor productivity					-0.2551			0.1285	
					(1.0309)			(0.9693)	
GDP per capita						-1.5635*			-3.3618***
						(0.8207)			(0.8952)
Spillovers automation							0.5442***	0.5478***	0.8587***
							(0.1831)	(0.1931)	(0.1270)
Spillovers other							-0.3014	-0.3089	-0.5853***
							(0.2573)	(0.2395)	(0.1790)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	50115	50115	50115	50115	50115	50115	50115	50115	50115
Firms	3341	3341	3341	3341	3341	3341	3341	3341	3341

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the country-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.18: Baseline regressions: effect of wage on automation innovations (auto90)

Dependent variable	Auto90								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	1.7307***	2.4414***	1.3357**	1.3715**	1.4738**	1.8797***	1.9059***	1.8309***	3.1623***
	(0.4953)	(0.6610)	(0.6363)	(0.6610)	(0.6778)	(0.7051)	(0.6883)	(0.7008)	(0.7486)
High-skill wage		-1.0613*	-0.7746	-0.8019	-0.6844	0.0911	-1.4074**	-1.5340**	-0.0865
		(0.5844)	(0.5311)	(0.5480)	(0.6068)	(0.5491)	(0.6296)	(0.6850)	(0.6114)
Stock automation			-0.0347	-0.0345	-0.0348	-0.0328	-0.0475	-0.0479	-0.0538
			(0.0405)	(0.0405)	(0.0404)	(0.0406)	(0.0403)	(0.0403)	(0.0403)
Stock other			0.5682***	0.5676***	0.5690***	0.5611***	0.5773***	0.5770***	0.5814***
			(0.0496)	(0.0497)	(0.0495)	(0.0495)	(0.0508)	(0.0508)	(0.0504)
GDP gap				0.0081	0.0085	0.0038	0.0152	0.0151	0.0127
				(0.0137)	(0.0134)	(0.0135)	(0.0133)	(0.0133)	(0.0132)
Labor productivity					-0.2904			0.2911	
					(0.7011)			(0.7224)	
GDP per capita						-2.0568***			-3.5341***
						(0.7380)			(0.7721)
Spillovers automation							0.8903**	0.9102**	1.2870***
							(0.4162)	(0.4190)	(0.4170)
Spillovers other							-0.6079**	-0.6342**	-1.0159***
							(0.3050)	(0.3140)	(0.3174)
Fixed Effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	73545	73545	73545	73545	73545	73545	73545	73545	73545
Firms	4903	4903	4903	4903	4903	4903	4903	4903	4903

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.19: Country-year fixed effects and country-level clustering

Dependent variable	Auto95								
	Domestic + Foreign			Foreign					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	1.8852**	2.1429***	3.0411***	3.4891***	4.3023**	3.7989**	3.6420***	4.3362**	3.8663**
	(0.8028)	(0.7524)	(1.1398)	(1.2222)	(1.9288)	(1.6359)	(1.3319)	(2.0053)	(1.5920)
High-skill wage	-2.4820***	-1.9117	-1.7526***	-3.5161**	-2.4740**	-3.3526***	-3.7549**	-2.8325***	-3.6398***
	(0.7416)	(1.3292)	(0.3511)	(1.5767)	(1.0274)	(1.2889)	(1.5240)	(0.9297)	(1.2942)
GDP gap	0.0623***	0.0620**	0.0646***	0.0044	0.0016	0.0044	0.0031	0.0001	0.0031
	(0.0239)	(0.0242)	(0.0216)	(0.0445)	(0.0397)	(0.0439)	(0.0456)	(0.0407)	(0.0452)
Labor productivity		-1.2851			-1.7494			-1.5475	
		(1.2933)			(1.6920)			(1.6342)	
GDP per capita			-2.8260			-0.5289			-0.3829
			(1.7682)			(1.3544)			(1.2045)
Stock automation	-0.1511***	-0.1506***	-0.1541***	-0.1522***	-0.1523***	-0.1526***	-0.1530***	-0.1532***	-0.1533***
	(0.0383)	(0.0382)	(0.0401)	(0.0371)	(0.0370)	(0.0373)	(0.0370)	(0.0370)	(0.0371)
Stock other	0.6549***	0.6556***	0.6555***	0.6494***	0.6471***	0.6490***	0.6496***	0.6475***	0.6493***
	(0.0532)	(0.0530)	(0.0543)	(0.0559)	(0.0570)	(0.0563)	(0.0555)	(0.0567)	(0.0559)
Spillovers automation	1.4782***	1.4762***	1.4715***	1.4396***	1.4128***	1.4355***	1.4380***	1.4161***	1.4357***
	(0.1276)	(0.1317)	(0.1188)	(0.1230)	(0.1585)	(0.1243)	(0.1243)	(0.1574)	(0.1254)
Spillovers other	-1.2259***	-1.2020***	-1.2436***	-1.2377***	-1.2268***	-1.2436***	-1.2252***	-1.2141***	-1.2300***
	(0.1690)	(0.1690)	(0.1633)	(0.1997)	(0.2111)	(0.1934)	(0.2002)	(0.2126)	(0.1941)
Fixed effects	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	50070	50070	50070	50070	50070	50070	50070	50070	50070
Firms	3338	3338	3338	3338	3338	3338	3338	3338	3338

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (4)-(6) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. In columns (7)-(9), they are interacted with the average shares over the sample period instead. In columns (4)-(9), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(3), there is no such interactions. Standard errors are clustered at the country-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.20: Country-year fixed effects and auto90

Dependent variable	Auto90								
	Domestic + Foreign			Foreign					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	1.3896*	1.4107	2.2798**	2.6344**	3.1221**	3.2536**	2.7215**	3.1094**	3.2428**
	(0.8386)	(0.8937)	(1.0390)	(1.1574)	(1.3170)	(1.3955)	(1.1927)	(1.3384)	(1.4122)
High-skill wage	-1.5576*	-1.5109	-1.0014	-3.0164**	-2.3531*	-2.6864**	-3.1666**	-2.6147*	-2.8915**
	(0.8304)	(0.9212)	(0.8793)	(1.2101)	(1.3149)	(1.2787)	(1.2485)	(1.3342)	(1.2984)
GDP gap	0.0387	0.0387	0.0405	-0.0044	-0.0060	-0.0042	-0.0053	-0.0070	-0.0053
	(0.0270)	(0.0270)	(0.0269)	(0.0361)	(0.0361)	(0.0360)	(0.0361)	(0.0362)	(0.0361)
Labor productivity		-0.1045			-1.0847			-0.8988	
		(1.1919)			(1.2059)			(1.1768)	
GDP per capita			-2.1599			-1.0595			-0.8978
			(1.4800)			(1.4139)			(1.3541)
Stock automation	-0.0537	-0.0536	-0.0556	-0.0572	-0.0576	-0.0577	-0.0577	-0.0580	-0.0581
	(0.0405)	(0.0406)	(0.0404)	(0.0405)	(0.0405)	(0.0405)	(0.0405)	(0.0404)	(0.0405)
Stock other	0.5846***	0.5847***	0.5845***	0.5802***	0.5794***	0.5792***	0.5802***	0.5796***	0.5795***
	(0.0510)	(0.0509)	(0.0508)	(0.0508)	(0.0507)	(0.0506)	(0.0508)	(0.0507)	(0.0506)
Spillovers automation	1.7794***	1.7789***	1.7682***	1.7676***	1.7438***	1.7562***	1.7652***	1.7459***	1.7563***
	(0.5417)	(0.5421)	(0.5434)	(0.5367)	(0.5388)	(0.5381)	(0.5357)	(0.5388)	(0.5370)
Spillovers other	-1.5492***	-1.5469***	-1.5563***	-1.5439***	-1.5316***	-1.5527***	-1.5350***	-1.5238***	-1.5431***
	(0.4359)	(0.4375)	(0.4366)	(0.4321)	(0.4320)	(0.4315)	(0.4305)	(0.4314)	(0.4298)
Fixed effects	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	73485	73485	73485	73485	73485	73485	73485	73485	73485
Firms	4899	4899	4899	4899	4899	4899	4899	4899	4899

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (4)-(6) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. In columns (7)-(9), they are interacted with the average shares over the sample period instead. In columns (4)-(9), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(3), there is no such interactions. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.21: Country-year dummies interacted with the domestic weight

Dependent variable	Auto95								
	Domestic + Foreign			Foreign					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	1.8108	2.3860*	2.2889*	2.0881*	2.6237**	2.9819**	2.1664*	2.6391**	2.9695**
	(1.1242)	(1.2486)	(1.3755)	(1.1178)	(1.2557)	(1.3805)	(1.1418)	(1.2624)	(1.3847)
High-skill wage	-2.7802**	-2.0793*	-2.5647**	-2.7271**	-2.1941*	-2.3615**	-2.9054**	-2.4236*	-2.5943**
	(1.1391)	(1.2117)	(1.1867)	(1.1229)	(1.2359)	(1.1984)	(1.1471)	(1.2481)	(1.2101)
GDP gap	0.0053	-0.0020	0.0021	0.0086	0.0037	0.0046	0.0075	0.0028	0.0039
	(0.0436)	(0.0444)	(0.0445)	(0.0440)	(0.0448)	(0.0445)	(0.0441)	(0.0449)	(0.0447)
Labor productivity		-1.2255			-0.9968			-0.9151	
		(0.9351)			(0.9758)			(0.9585)	
GDP per capita			-0.7515			-1.3618			-1.2168
			(1.2918)			(1.3924)			(1.3560)
Stock automation	-0.1531***	-0.1525***	-0.1531***	-0.1518***	-0.1514***	-0.1523***	-0.1519***	-0.1515***	-0.1525***
	(0.0523)	(0.0521)	(0.0522)	(0.0522)	(0.0520)	(0.0521)	(0.0522)	(0.0520)	(0.0520)
Stock other	0.6433***	0.6417***	0.6429***	0.6420***	0.6407***	0.6412***	0.6422***	0.6409***	0.6415***
	(0.0605)	(0.0603)	(0.0603)	(0.0607)	(0.0606)	(0.0603)	(0.0607)	(0.0606)	(0.0603)
Spillovers automation	1.1705***	1.2209***	1.2079***	1.0883**	1.1219***	1.1442***	1.1121***	1.1484***	1.1663***
	(0.4154)	(0.4139)	(0.4199)	(0.4241)	(0.4227)	(0.4283)	(0.4191)	(0.4183)	(0.4241)
Spillovers other	-0.9536***	-0.9457***	-0.9736***	-0.9431***	-0.9441***	-0.9801***	-0.9379***	-0.9386***	-0.9719***
	(0.3302)	(0.3305)	(0.3319)	(0.3315)	(0.3310)	(0.3333)	(0.3315)	(0.3315)	(0.3335)
Fixed effects	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	50085	50085	50085	50085	50085	50085	50085	50085	50085
Firms	3339	3339	3339	3339	3339	3339	3339	3339	3339

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm and country-year fixed effects. Country-year fixed effects are interacting with the countries' weights. All regressions with stock variables include a dummy for no stock and no spillover. In columns (4)-(6) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. In columns (7)-(9), they are interacted with the average shares over the sample period instead. In columns (4)-(9), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(3), there is no such interactions. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.22: All machinery innovations

Dependent variable	Machinery								
	Domestic + Foreign						Foreign		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	0.4615	0.5068	1.7568***	0.1366	0.1813	1.1875	-0.0553	0.4116	0.9712
	(0.5070)	(0.5585)	(0.5336)	(0.6721)	(0.7338)	(0.8614)	(1.0951)	(1.2570)	(1.2306)
High-skill wage	-0.0290	0.0429	1.2020**	-0.0638	0.0273	0.5109	-0.2290	0.3470	0.2678
	(0.5224)	(0.4950)	(0.5625)	(0.7663)	(0.8065)	(0.7838)	(1.2089)	(1.1884)	(1.2571)
GDP gap	-0.0211**	-0.0209**	-0.0219**	0.0080	0.0080	0.0100	0.0228	0.0221	0.0239
	(0.0086)	(0.0084)	(0.0086)	(0.0153)	(0.0153)	(0.0153)	(0.0235)	(0.0235)	(0.0234)
Labor productivity		-0.1676			-0.1986			-0.9600	
		(0.6030)			(0.9082)			(0.9293)	
GDP per capita			-3.5955***			-2.3745**			-1.6536
			(0.6309)			(1.1561)			(1.0728)
Stock machinery	0.3337***	0.3339***	0.3337***	0.3400***	0.3401***	0.3386***	0.3379***	0.3372***	0.3370***
	(0.0352)	(0.0352)	(0.0341)	(0.0341)	(0.0342)	(0.0340)	(0.0345)	(0.0346)	(0.0345)
Stock other	0.2443***	0.2444***	0.2416***	0.2456***	0.2458***	0.2454***	0.2449***	0.2442***	0.2435***
	(0.0441)	(0.0442)	(0.0426)	(0.0424)	(0.0423)	(0.0420)	(0.0413)	(0.0414)	(0.0413)
Spillovers machinery	2.7148***	2.7216***	1.9670***	1.1095**	1.1138**	1.0117**	1.0863**	1.1211**	1.0510**
	(0.4332)	(0.4335)	(0.4422)	(0.4652)	(0.4692)	(0.4640)	(0.4638)	(0.4666)	(0.4644)
Spillovers other	-2.4318***	-2.4309***	-1.8095***	-1.2125**	-1.2136**	-1.1326**	-1.1784**	-1.2191**	-1.1742**
	(0.5096)	(0.5110)	(0.5022)	(0.4859)	(0.4867)	(0.4857)	(0.4802)	(0.4839)	(0.4801)
Fixed effects	F + Y	F + Y	F + Y	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	160290	160290	160290	160290	160290	160290	160290	160290	160290
Firms	10686	10686	10686	10686	10686	10686	10686	10686	10686

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). In columns (1)-(3), the regressions include firm and year fixed effects. In columns (4)-(9), the regressions include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (7)-(9) domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. In columns (7)-(9), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. In columns (1)-(6), there is no such interactions. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.23: Predicted wages

Dependent Variable	Auto95							
	joint ρ, average		joint ρ, t+4		separate ρ, average		separate ρ, t+4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low-skill wage	1.6899**	1.4813*	1.7039**	1.4899*	1.7557**	1.4318*	1.7803**	1.4313*
	(0.8152)	(0.8080)	(0.8167)	(0.8107)	(0.8286)	(0.8087)	(0.8314)	(0.8137)
High-skill wage	-1.7960**	-2.9855**	-1.7638**	-2.8597*	-1.7838**	-2.7068**	-1.7874**	-2.7378**
	(0.8440)	(1.5046)	(0.8440)	(1.4860)	(0.8196)	(1.2652)	(0.8283)	(1.2776)
GDP gap	0.0162	0.0161	0.0164	0.0163	0.0144	0.0119	0.0144	0.0117
	(0.0143)	(0.0143)	(0.0143)	(0.0143)	(0.0142)	(0.0139)	(0.0142)	(0.0139)
Labor productivity		1.7353		1.6234		1.4848		1.5467
		(1.7310)		(1.7208)		(1.1824)		(1.2247)
Stock automation	-0.1433***	-0.1451***	-0.1430***	-0.1446***	-0.1433***	-0.1451***	-0.1431***	-0.1450***
	(0.0509)	(0.0514)	(0.0509)	(0.0514)	(0.0509)	(0.0517)	(0.0510)	(0.0517)
Stock other	0.6408***	0.6380***	0.6407***	0.6379***	0.6405***	0.6371***	0.6405***	0.6371***
	(0.0601)	(0.0603)	(0.0601)	(0.0603)	(0.0602)	(0.0604)	(0.0601)	(0.0604)
Spillovers automation	0.4847	0.6321*	0.4848	0.6209*	0.5049*	0.7348**	0.5097*	0.7364**
	(0.3045)	(0.3449)	(0.3049)	(0.3445)	(0.3036)	(0.3702)	(0.3044)	(0.3679)
Spillovers other	-0.1628	-0.3290	-0.1674	-0.3214	-0.1842	-0.4488	-0.1899	-0.4498
	(0.2276)	(0.2877)	(0.2278)	(0.2866)	(0.2281)	(0.3182)	(0.2282)	(0.3152)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	50115	50115	50115	50115	50115	50115	50115	50115
Firms	3341	3341	3341	3341	3341	3341	3341	3341

Note: Marginal effects; Standard errors in parentheses. Estimation is by conditional Poisson regressions fixed-effects (HHG). The wage variables and labor productivity are predicted at time t-2. Columns (1) to (4) predict wages and labor productivity with an AR(1) process with country-specific trends and with the same auto-regression coefficient across countries. Columns (5) to (8) use different auto-regression coefficients across countries. In columns (1), (2), (5) and (6) the wages and labor productivity are the average of the predicted values between years t+2 and t+7. In columns (3), (4), (7) and (8), they are the predicted values for year t+4. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.24: Minimum wage

Dependent variable	Auto95							
	Domestic + Foreign				Foreign			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Minimum wage	1.5230**	1.5171**	1.4636	1.5601	1.8773	1.8401	1.8331*	1.7770
	(0.6865)	(0.6628)	(0.9127)	(0.9566)	(1.2125)	(1.2411)	(1.1117)	(1.1271)
High-skill wage	-1.2239*	-1.2358	-3.0712***	-2.6564**	-2.8017**	-2.9368	-2.7456**	-3.0195*
	(0.7166)	(0.8701)	(1.0907)	(1.1667)	(1.4072)	(1.8000)	(1.3157)	(1.7324)
GDP gap	0.0235	0.0235	0.0562	0.0563	-0.0232	-0.0232	-0.0238	-0.0236
	(0.0151)	(0.0150)	(0.0347)	(0.0347)	(0.0513)	(0.0514)	(0.0512)	(0.0513)
Labor productivity		0.0246		-0.7554		0.1730		0.3355
		(0.9249)		(1.4016)		(1.4426)		(1.3849)
Stock automation	-0.1445***	-0.1446***	-0.1548***	-0.1544***	-0.1563***	-0.1564***	-0.1562***	-0.1565***
	(0.0513)	(0.0513)	(0.0522)	(0.0523)	(0.0530)	(0.0531)	(0.0530)	(0.0531)
Stock other	0.6374***	0.6374***	0.6569***	0.6572***	0.6549***	0.6552***	0.6540***	0.6547***
	(0.0596)	(0.0596)	(0.0597)	(0.0597)	(0.0607)	(0.0607)	(0.0607)	(0.0606)
Spillovers automation	0.6456*	0.6462*	1.4309***	1.4270***	1.4198***	1.4215***	1.4157***	1.4192***
	(0.3363)	(0.3397)	(0.4958)	(0.4967)	(0.4939)	(0.4966)	(0.4929)	(0.4955)
Spillovers other	-0.3546	-0.3559	-1.1991***	-1.1837***	-1.2744***	-1.2764***	-1.2806***	-1.2846***
	(0.2408)	(0.2535)	(0.3854)	(0.3864)	(0.3795)	(0.3821)	(0.3787)	(0.3810)
Fixed effects	F	F	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	50070	50070	50040	50040	48765	48765	48795	48795
Firms	3338	3338	3336	3336	3251	3251	3254	3254

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Columns (1)-(2) include firm fixed effects. Columns (3)-(8) include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (5)-(6) domestic (resp. foreign) minimum wages are interacted with the share of domestic (resp. foreign) minimum wages in total minimum wages computed at the beginning of the sample, and similarly for high-skill wages and VA per employee. In columns (7)-(8), they are interacted with the average shares over the sample period instead. In columns (5)-(8), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.25: Minimum wage with country level clustering

Dependent variable	Auto95							
	Domestic + Foreign				Foreign			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Minimum wage	1.5230**	1.5171***	1.4636**	1.5601***	1.8773***	1.8401***	1.8331***	1.7770***
	(0.5926)	(0.4580)	(0.6148)	(0.4905)	(0.4685)	(0.6459)	(0.4272)	(0.5790)
High-skill wage	-1.2239**	-1.2358	-3.0712***	-2.6564**	-2.8017***	-2.9368***	-2.7456***	-3.0195***
	(0.5538)	(1.0144)	(0.5048)	(1.2687)	(1.0073)	(0.8003)	(0.9635)	(0.8786)
GDP gap	0.0235***	0.0235***	0.0562***	0.0563***	-0.0232	-0.0232	-0.0238	-0.0236
	(0.0046)	(0.0047)	(0.0209)	(0.0210)	(0.0246)	(0.0245)	(0.0238)	(0.0235)
Labor productivity		0.0246		-0.7554		0.1730		0.3355
		(0.9997)		(1.4283)		(1.3091)		(1.3969)
Stock automation	-0.1445***	-0.1446***	-0.1548***	-0.1544***	-0.1563***	-0.1564***	-0.1562***	-0.1565***
	(0.0385)	(0.0390)	(0.0403)	(0.0400)	(0.0392)	(0.0402)	(0.0391)	(0.0404)
Stock other	0.6374***	0.6374***	0.6569***	0.6572***	0.6549***	0.6552***	0.6540***	0.6547***
	(0.0514)	(0.0513)	(0.0563)	(0.0561)	(0.0572)	(0.0595)	(0.0569)	(0.0598)
Spillovers automation	0.6456***	0.6462***	1.4309***	1.4270***	1.4198***	1.4215***	1.4157***	1.4192***
	(0.2076)	(0.2225)	(0.1139)	(0.1151)	(0.1192)	(0.1309)	(0.1182)	(0.1314)
Spillovers other	-0.3546	-0.3559	-1.1991***	-1.1837***	-1.2744***	-1.2764***	-1.2806***	-1.2846***
	(0.2214)	(0.2362)	(0.1684)	(0.1736)	(0.1956)	(0.2102)	(0.1920)	(0.2072)
Fixed effects	F	F	F + CY	F + CY	F + CY	F + CY	F + CY	F + CY
Observations	50070	50070	50040	50040	48765	48765	48795	48795
Firms	3338	3338	3336	3336	3251	3251	3254	3254

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Columns (1)-(2) include firm fixed effects. Columns (3)-(8) include firm and country-year fixed effects. All regressions with stock variables include a dummy for no stock and no spillover. In columns (5)-(6) domestic (resp. foreign) minimum wages are interacted with the share of domestic (resp. foreign) minimum wages in total minimum wages computed at the beginning of the sample, and similarly for high-skill wages and VA per employee. In columns (7)-(8), they are interacted with the average shares over the sample period instead. In columns (5)-(8), domestic (resp. foreign) GDP gap is interacted with the domestic (resp. foreign) weight. Standard errors are clustered at the country-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.26: Correlation matrix

	Low-skill wage	Middle-skill wage	High-skill wage	GDP gap	GDP per capita	Labor productivity
Low-skill wage	1	.	.	.	.	.
Middle-skill wage	0.9401	1	.	.	.	.
High-skill wage	0.6009	0.7469	1		.	.
GDP gap	-0.0660	-0.0239	0.0482	1		.
GDP per capita	0.6972	0.7974	0.7277	-0.0117	1	.
Labor productivity	0.6678	0.7340	0.7724	0.1980	0.6519	1

Note: Correlation of residuals for the auto95 sample controlling for year and firm fixed effects.

Table 1.8.27: Skill premium

Dependent variable	Auto95					
	Domestic + Foreign				Foreign	
	(1)	(2)	(3)	(4)	(5)	(6)
Low-skill / High-skill wage	1.9423**	2.0420***	2.1995**	2.0520**	3.5089***	3.4205***
	(0.7552)	(0.7607)	(0.9170)	(0.9049)	(1.2083)	(1.1960)
GDP gap	0.0263*	0.0268*	0.0627*	0.0620*	0.0049	-0.0017
	(0.0157)	(0.0157)	(0.0343)	(0.0343)	(0.0526)	(0.0496)
Labor productivity		0.7026		-1.0613		-0.2814
		(0.7035)		(1.1591)		(0.7369)
Stock automation	-0.1448***	-0.1456***	-0.1505***	-0.1507***	-0.1522***	-0.1524***
	(0.0509)	(0.0510)	(0.0530)	(0.0528)	(0.0526)	(0.0525)
Stock other	0.6407***	0.6402***	0.6546***	0.6556***	0.6495***	0.6480***
	(0.0599)	(0.0601)	(0.0603)	(0.0602)	(0.0602)	(0.0602)
Spillovers automation	0.5783*	0.5783*	1.4755***	1.4769***	1.4397***	1.4346***
	(0.3153)	(0.3114)	(0.4968)	(0.5004)	(0.4868)	(0.4892)
Spillovers other	-0.2349	-0.3132	-1.2535***	-1.2021***	-1.2387***	-1.2253***
	(0.2129)	(0.2328)	(0.3717)	(0.3824)	(0.3669)	(0.3755)
Fixed effects	F + Y	F + Y	F + CY	F + CY	F + CY	F + CY
Observations	50115	50115	50070	50070	50070	50070
Firms	3341	3341	3338	3338	3338	3338

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). Columns (1)-(2) include firm fixed effects and year dummies. Columns (3)-(6) include firm and country-year fixed effects. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Columns (5)-(6) use the log difference between foreign low-skill wages interacted with the share of foreign low-skill wages in total low-skill wages at the beginning of the sample and foreign high-skill wages similarly interacted; GDP gap and VA per employee are also their interacted foreign components. Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.28: Including middle-skill wages

Dependent Variable	Auto95								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	4.7035***		3.8985***	5.1140***		4.2760***	5.0971***		4.2398***
	(1.4991)		(1.3667)	(1.5892)		(1.4222)	(1.5759)		(1.4510)
Middle-skill wage	-3.9194**	2.3617**	-2.2614	-4.2997**	2.4746**	-2.5516	-4.0739**	2.3236**	-2.5526
	(1.6096)	(1.0085)	(1.6773)	(1.6815)	(1.0411)	(1.6819)	(1.6090)	(1.0573)	(1.6825)
High-skill wage		-1.7189*	-0.9608		-1.8154*	-1.0225		-1.9749**	-1.0756
		(0.9218)	(0.8867)		(0.9485)	(0.8960)		(0.9982)	(0.9602)
GDP gap				0.0288*	0.0216	0.0304*	0.0292*	0.0216	0.0303*
				(0.0153)	(0.0151)	(0.0157)	(0.0154)	(0.0151)	(0.0157)
Labor productivity							-0.3258	0.4545	0.1316
							(0.8572)	(0.8858)	(0.9310)
Stock automation	-0.1454***	-0.1404***	-0.1457***	-0.1460***	-0.1405***	-0.1464***	-0.1455***	-0.1415***	-0.1466***
	(0.0508)	(0.0508)	(0.0509)	(0.0509)	(0.0509)	(0.0510)	(0.0509)	(0.0510)	(0.0511)
Stock other	0.6458***	0.6394***	0.6436***	0.6456***	0.6389***	0.6433***	0.6455***	0.6387***	0.6432***
	(0.0598)	(0.0598)	(0.0600)	(0.0599)	(0.0600)	(0.0601)	(0.0599)	(0.0600)	(0.0601)
Spillovers automation	0.4733	0.4518	0.5330*	0.5007*	0.4692	0.5657*	0.4998*	0.4846	0.5694*
	(0.2891)	(0.3140)	(0.3097)	(0.2885)	(0.3143)	(0.3105)	(0.2891)	(0.3155)	(0.3120)
Spillovers other	-0.3173	-0.1874	-0.3100	-0.3478	-0.2013	-0.3416	-0.3281	-0.2315	-0.3492
	(0.2254)	(0.2208)	(0.2265)	(0.2247)	(0.2197)	(0.2257)	(0.2315)	(0.2279)	(0.2325)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	50115	50115	50115	50115	50115	50115	50115	50115	50115
Firms	3341	3341	3341	3341	3341	3341	3341	3341	3341

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions with stock variables include a dummy for no stock and no spillover. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.29: Wages and deflators

Dependent variable	Auto95									
Sector	Manufacturing						Total			
Deflator	Manufacturing PPI, conversion in 2005		US manufacturing PPI conversion every year		GDP deflator conversion in 1995		Manufacturing PPI conversion in 1995		US Manufacturing PPI conversion every year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Low-skill wage	2.7140***	2.6338***	1.9084***	2.1264**	2.5733***	2.7044***	4.1859***	3.9769***	1.4172**	1.1137
	(0.8686)	(0.8933)	(0.6949)	(0.8261)	(0.9691)	(1.0238)	(1.3286)	(1.2666)	(0.7192)	(0.8024)
High-skill wage	-1.7475**	-1.8694**	-2.4692***	-2.2154***	-2.1163**	-1.9409**	-1.3163	-2.3907**	-2.0329***	-2.3743**
	(0.7943)	(0.8603)	(0.7517)	(0.7790)	(0.9229)	(0.9578)	(0.8454)	(0.9545)	(0.7025)	(0.9521)
GDP gap	0.0285*	0.0283*	0.0153	0.0149	0.0254	0.0262	0.0431**	0.0440**	0.0158	0.0148
	(0.0158)	(0.0158)	(0.0146)	(0.0146)	(0.0161)	(0.0161)	(0.0171)	(0.0172)	(0.0152)	(0.0153)
Labor productivity		0.3056		-0.5012		-0.4125		2.6369**		0.6389
		(0.9422)		(0.7122)		(0.7779)		(1.2281)		(0.9348)
Stock own	-0.1439***	-0.1444***	-0.1501***	-0.1493***	-0.1454***	-0.1446***	-0.1446***	-0.1474***	-0.1457***	-0.1462***
	(0.0510)	(0.0511)	(0.0510)	(0.0510)	(0.0510)	(0.0511)	(0.0506)	(0.0509)	(0.0506)	(0.0508)
Stock other	0.6392***	0.6390***	0.6391***	0.6396***	0.6403***	0.6405***	0.6485***	0.6455***	0.6434***	0.6424***
	(0.0600)	(0.0601)	(0.0598)	(0.0597)	(0.0600)	(0.0599)	(0.0596)	(0.0596)	(0.0592)	(0.0592)
Spillover own	0.5795*	0.5887*	0.8540**	0.8568**	0.6503*	0.6444*	0.4874*	0.5675**	0.6379**	0.6536**
	(0.3073)	(0.3093)	(0.3471)	(0.3459)	(0.3451)	(0.3456)	(0.2862)	(0.2879)	(0.3217)	(0.3275)
Spillover other	-0.3314	-0.3499	-0.4295*	-0.4312*	-0.3447	-0.3310	-0.2943	-0.4228*	-0.2826	-0.2962
	(0.2259)	(0.2344)	(0.2332)	(0.2332)	(0.2220)	(0.2219)	(0.2399)	(0.2510)	(0.2403)	(0.2414)
Fixed Effect	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	50115	50115	50115	50115	50115	50115	50115	50115	50115	50115
Firms	3341	3341	3341	3341	3341	3341	3341	3341	3341	3341
Clustering	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions include a dummy for no stock and no spillover. Columns (1) to (6) are on manufacturing wages and columns (7) to (10) on total wages. In columns (1) and (2), macroeconomic variables are deflated with the local manufacturing PPI and converted in USD in 2005. In columns (3), (4), (9) and (10) they are converted in USD every year and deflated with the US manufacturing PPI. In columns (5) and (6), macroeconomic variables are deflated with the local GDP deflator and converted in USD in 1995. In columns (7) and (8), macroeconomic variables are deflated with the local manufacturing PPI and converted in USD in 1995. Standard errors are clustered at the firm-level * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.30: Baseline regressions in 2000-20009 only

Dependent variable	Auto95								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	2.6434***	2.4828**	1.8409*	1.9912*	2.0772*	2.4664**	2.9215**	2.6339**	4.4721***
	(0.7284)	(0.9935)	(1.0261)	(1.0700)	(1.0718)	(1.2056)	(1.1447)	(1.1285)	(1.4064)
High-skill wage		0.2690	-0.3563	-0.5113	-0.4602	-0.2960	-1.1540	-1.4516	-0.7074
		(0.8835)	(0.8904)	(0.9219)	(0.9557)	(0.8762)	(0.9894)	(1.0716)	(0.9416)
Stock automation			-0.4117***	-0.4100***	-0.4105***	-0.4050***	-0.4375***	-0.4398***	-0.4335***
			(0.0630)	(0.0631)	(0.0628)	(0.0635)	(0.0636)	(0.0639)	(0.0639)
Stock other			0.6746***	0.6708***	0.6725***	0.6687***	0.6881***	0.6864***	0.6937***
			(0.0709)	(0.0711)	(0.0714)	(0.0708)	(0.0744)	(0.0743)	(0.0735)
GDP gap				0.0243	0.0246	0.0196	0.0419**	0.0437**	0.0360**
				(0.0164)	(0.0162)	(0.0157)	(0.0171)	(0.0174)	(0.0169)
Labor productivity					-0.1968			1.1082	
					(0.9325)			(0.9940)	
GDP per capita						-1.5031			-3.7815**
						(1.1155)			(1.4968)
Spillovers automation							0.9119**	1.0198**	1.1483***
							(0.4167)	(0.4249)	(0.4267)
Spillovers other							-0.5948*	-0.7380*	-0.8383**
							(0.3577)	(0.3820)	(0.3731)
Fixed effects	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y	F + Y
Observations	27110	27110	27110	27110	27110	27110	27110	27110	27110
Firms	2711	2711	2711	2711	2711	2711	2711	2711	2711

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG) from 2000 to 2009. All regressions include firm fixed effects and year dummies. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.31: Nickell's bias

Dependent Variable	Auto95							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low-skill wage	2.3903***	2.3926***	2.1515***	2.2066***	2.0925**	2.2884**	2.3955**	2.9126***
	(0.8004)	(0.8227)	(0.7991)	(0.8150)	(0.9778)	(1.0886)	(0.9713)	(1.0899)
High-skill wage	-1.5544**	-1.5510*	-0.9069	-0.5857	-2.4648**	-2.0312**	-2.5627***	-1.2324
	(0.7840)	(0.8704)	(0.6129)	(0.7453)	(0.9779)	(0.9708)	(0.9338)	(1.0583)
GDP gap	0.0276*	0.0276*	0.0266	0.0278	0.0653*	0.0651*	0.0752**	0.0761**
	(0.0159)	(0.0158)	(0.0191)	(0.0187)	(0.0343)	(0.0342)	(0.0353)	(0.0353)
Labor productivity		-0.0084		-0.7779		-0.9781		-2.6421
		(0.9696)		(1.0755)		(1.5602)		(1.6507)
Stock automation			1.1938***	1.1818***			1.1912***	1.1870***
			(0.0244)	(0.0238)			(0.0243)	(0.0235)
Stock other	0.5101***	0.5101***	0.0895***	0.0897***	0.5230***	0.5237***	0.0869***	0.0879***
	(0.0454)	(0.0453)	(0.0120)	(0.0118)	(0.0439)	(0.0440)	(0.0120)	(0.0118)
Spillovers automation	0.3519	0.3517	0.0098	-0.0315	1.3383***	1.3373***	-0.0667	-0.0518
	(0.2949)	(0.2977)	(0.0746)	(0.0689)	(0.4669)	(0.4676)	(0.0784)	(0.0767)
Spillovers other	-0.0735	-0.0730	0.0219	0.0781	-1.0318***	-1.0139***	0.1163	0.1013
	(0.2127)	(0.2227)	(0.0782)	(0.0748)	(0.3544)	(0.3558)	(0.0827)	(0.0815)
Fixed effects	F + Y	F + Y	BGVR + Y	BGVR + Y	F + CY	F + CY	BGVR + CY	BGVR + CY
Observations	50115	50115	50115	50115	50070	50070	50070	50070
Firms	3341	3341	3341	3341	3338	3338	3338	3338

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG) in columns (1), (2), (5) and (6). In columns (3), (4), (7) and (8), estimation is done by Poisson regressions where the firm fixed effects are replaced by the pre-sample mean, following Blundell, Griffith and Van Reenen (1999, BGVR). Columns (1) to (4) include year fixed effects and columns (5) to (8) country-year fixed effects. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.32: Other innovation indicators

	Auto95							
Dependent Variable	Biadic (US, JP, EU)		Triadic		At least one citation		Citations weighted	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low-skill wage	2.2776**	2.0079*	3.1886**	2.9795*	2.2198***	2.1241**	1.7405*	1.6520
	(1.0383)	(1.0785)	(1.4150)	(1.5827)	(0.8341)	(0.8720)	(1.0257)	(1.1403)
High-skill wage	-1.3409	-1.7718*	-2.3417*	-2.6759*	-1.6034**	-1.7443**	-1.8007*	-1.9515**
	(0.9663)	(1.0724)	(1.3640)	(1.3768)	(0.8099)	(0.8577)	(0.9814)	(0.9717)
GDP gap	0.0397**	0.0390**	0.0178	0.0172	0.0269*	0.0267*	0.0368*	0.0366*
	(0.0191)	(0.0191)	(0.0289)	(0.0288)	(0.0158)	(0.0157)	(0.0190)	(0.0190)
Labor productivity		0.9807		0.7272		0.3450		0.3518
		(1.1988)		(1.6987)		(0.9171)		(1.1755)
Stock automation	-0.1683***	-0.1699***	-0.3665***	-0.3677***	-0.1468***	-0.1474***	-0.2220***	-0.2223***
	(0.0597)	(0.0598)	(0.0772)	(0.0766)	(0.0557)	(0.0559)	(0.0438)	(0.0438)
Stock other	0.6342***	0.6333***	0.6500***	0.6494***	0.6457***	0.6456***	0.6805***	0.6802***
	(0.0662)	(0.0663)	(0.0875)	(0.0875)	(0.0635)	(0.0635)	(0.0688)	(0.0687)
Spillovers automation	0.3839	0.4064	0.7925	0.7981	0.5736*	0.5845*	0.1427	0.1499
	(0.4014)	(0.4028)	(0.5469)	(0.5451)	(0.3140)	(0.3151)	(0.2878)	(0.2858)
Spillovers other	-0.5402**	-0.5915**	-0.3499	-0.3742	-0.2978	-0.3187	0.1625	0.1429
	(0.2587)	(0.2715)	(0.4685)	(0.4599)	(0.2404)	(0.2468)	(0.2595)	(0.2600)
Observations	40410	40410	26310	26310	47115	47115	50115	50115
Firms	2694	2694	1754	1754	3141	3141	3341	3341

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditiona Poisson regressions fixed-effects (HHG). All regressions include firm fixed effects and year dummies. All regressions include a dummy for nc stock and no spillover. Columns (1)-(2) consider biadic patents applied for in at least two countries among US, JP, EU. Columns (3)-(4 consider triadic patents (applied for in US, JP and EU). Column (5)-(6) consider biadic patents with at least one citation within 5 year after publication. Column (7)-(8) consider biadic patents and add to each patent the number of citations within 5 years after publicatio normalized by year and technological field. Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1.8.33: Industry-year fixed effects

Dependent variable	Auto95								
	Domestic + Foreign			Foreign					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low-skill wage	2.3157**	2.5169**	3.5773***	4.1573***	5.0264***	4.2013**	4.3562***	5.0852***	4.3077**
	(0.9890)	(1.1159)	(1.2188)	(1.3041)	(1.5426)	(1.7227)	(1.3302)	(1.5430)	(1.7067)
High-skill wage	-2.9978***	-2.5654**	-2.1617**	-4.3227***	-3.1470**	-4.2974***	-4.5869***	-3.5601**	-4.6144***
	(0.9457)	(1.0210)	(1.0263)	(1.2915)	(1.3761)	(1.3321)	(1.3283)	(1.3944)	(1.3495)
GDP gap	0.0709**	0.0707**	0.0731**	-0.0059	-0.0083	-0.0059	-0.0066	-0.0093	-0.0065
	(0.0323)	(0.0323)	(0.0322)	(0.0470)	(0.0469)	(0.0470)	(0.0471)	(0.0470)	(0.0471)
Labor productivity		-0.9736			-1.9354			-1.6865	
		(1.7031)			(1.4734)			(1.4339)	
GDP per capita			-3.1161*			-0.0777			0.0860
			(1.7989)			(1.8617)			(1.7683)
Stock automation	-0.1586***	-0.1582***	-0.1601***	-0.1607***	-0.1603***	-0.1607***	-0.1617***	-0.1615***	-0.1617***
	(0.0466)	(0.0466)	(0.0468)	(0.0463)	(0.0461)	(0.0464)	(0.0462)	(0.0460)	(0.0463)
Stock other	0.6549***	0.6555***	0.6548***	0.6492***	0.6470***	0.6491***	0.6497***	0.6478***	0.6497***
	(0.0552)	(0.0552)	(0.0549)	(0.0550)	(0.0549)	(0.0549)	(0.0549)	(0.0548)	(0.0548)
Spillovers automation	1.3924***	1.3897***	1.3786***	1.3568***	1.3324***	1.3562***	1.3558***	1.3371***	1.3563***
	(0.4759)	(0.4766)	(0.4761)	(0.4658)	(0.4651)	(0.4675)	(0.4653)	(0.4651)	(0.4665)
Spillovers other	-1.0750***	-1.0587***	-1.0900***	-1.0864***	-1.0802***	-1.0873***	-1.0740***	-1.0670***	-1.0730***
	(0.3623)	(0.3642)	(0.3618)	(0.3553)	(0.3527)	(0.3540)	(0.3538)	(0.3520)	(0.3527)
Fixed effects	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY	F + CY +IY
Observations	49890	49890	49890	49890	49890	49890	49890	49890	49890
Firms	3326	3326	3326	3326	3326	3326	3326	3326	3326

Note: Marginal effects; Standard errors in parentheses. The independent variables are lagged by two periods. Estimation is by conditional Poisson regressions fixed-effects (HHG). All regressions include firm, industry-year and country-year fixed effects. All regressions with stock variables (resp. spillover variables) include a dummy for no stock (resp. no spillover). Domestic (resp. foreign) low-skill wages are interacted with the share of domestic (resp. foreign) low-skill wages in total low-skill wages computed at the beginning of the sample, and similarly for high-skill wages, GDP per capita and VA per employee. Standard errors are clustered at the firm-level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

1.8.2 Details on the classification of automation patents

List of keywords

For each technological category, we compute the following share of patents:[41]

1. Automat* patents. Share of patents which contain the words:

   (a) *Automation* or *automatization;*

   (b) or *automat** at least 5 times;

   (c) or (*automat** or *autonomous*) in the same sentence as (*machine* or *manufacturing* or *machining* or *equipment* or *apparatus* or *operator* or *handling* or *"vehicle system"* or *welding* or *knitting* or *weaving* or *convey** or *storage* or *store* or *regulat** or *manipulat** or *arm* or *sensor* or *inspect** or *warehouse*) at least twice.

2. Labor patents. Share of patents which contain the words: *laborious*, *labourious, labor* or *labour.*

[41] x* indicates any word which starts with x, for instance automat* corresponds to the words automatic, automatically, automate, automates, etc...

3. Robot patents. Share of patents which contain the word *robot** but not (*surgical* or *medical*).

4. Numerical control patents. Share of patents which contain the words:

   (a) *CNC* or "*numerically controlled*" or "*numeric control*" or "*numerical control*" or the same terms but with hyphens;

   (b) or *NC* in the same sentence with (*machine* or *manufacturing* or *machining* or *equipment* or *apparatus*).

5. Computer aided design and manufacturing patents. Share of patents which contain the words:

   (a) "*computer aided*", "*computer assisted*" or "*computer supported*" or the same terms with hyphens) in the same patent with (*machine* or *manufacturing* or *machining* or *equipment* or *apparatus*);

   (b) or (*CAD* or (*CAM* and not "*content addressable memory*")) in the same sentence with (*machine* or *manufacturing* or *machining* or *equipment* or *apparatus*).

6. Flexible manufacturing. Share of patents which contain the words: "*flexible manufacturing*".

7. PLC patents. Share of patents which contain the words: "*programmable logic controller*" or (*PLC* and not (*powerline* or "*power line*")).

8. 3D printing patents. Share of patents which contain the words: "*3D print**" or "*additive manufacturing*" or "*additive layer manufacturing*".

9. Automation patents. Share of patents which satisfy any of the previous criteria.

We derived this exact list after experimenting extensively with variations around those words and looking at the resulting classification of technological codes and the associated patents. For instance, the thresholds (5 and 2) used in the definition of the share of automat* patents where chosen so that the distribution of the share of automat* patents is comparable to the distribution of the share of numerical control patents across technological codes. Similarly, requiring that *NC* be in the same sentence as words such as *machine*, ensures that *NC* is short fort numerical control instead of North Carolina.

Relative to the original list of technologies given in the SMT, we did not include keywords related to information network, as these seem less related to the automation of the production process and the patents containing words such as "local area network" do not appear related to automation. We also did not directly count all laser related technologies as not all of these are related to automation—but we obtain patents related to automation using laser technologies thanks to our other keywords.

Statistics on the classification

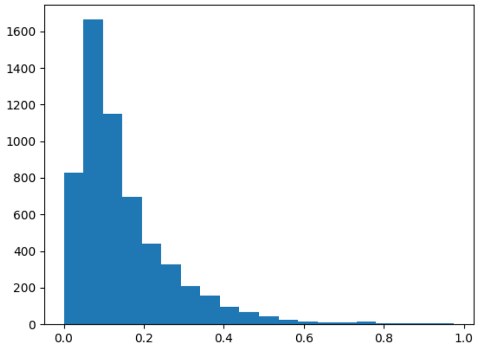
Table 1.8.34: Summary statistics on the prevalence of keywords across technological codes in machinery

	IPC/CPC 6 digit				IPC4 + (G05 or G06)				IPC 4 pairs			
Share	all	robot	automat*	CNC	all	robot	automat*	CNC	all	robot	automat*	CNC
Mean	20.9	4.3	11.2	2.4	53.2	15.4	32.4	11.2	18.5	4.5	8.8	1.8
S. d.	14.4	8.4	9.5	5.8	19.3	17.7	11	16.5	16.3	10	9.9	4.7
p25	10.5	0.8	4.2	0	40	6.7	26.6	0.8	7.7	0.6	2.5	0
p50	18	2	8.7	0.4	54.3	10	31.9	3	13.6	1.8	5.2	0.4
p75	26.6	4.5	15.3	1.8	63.8	16	40.3	15.5	23	4.2	10.7	1.4
p90	38.7	9.1	24.3	6.1	77.9	36.4	43.3	38.2	36.8	8.9	21.7	4.4
p95	47.7	13.7	29.4	12.7	85.6	44.3	45.2	55.3	51.8	14.5	31	7.7
p99	75	35.8	43.8	33.1	90.1	82.9	59.9	56.6	84.5	60	45.3	23.1

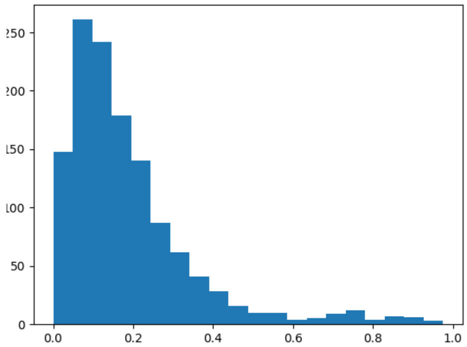
Note: This table computes summary statistics on the share of patents with any automation keywords, robot keywords, automat* keywords or CNC keywords for each type of technological categories (6 digit codes, pairs of 4 digit codes and combinations of ipc4 codes with G05 or G06) within machinery with at least 100 patents.

Table 1.8.34 gives summary statistics on the shares of patents containing certain keywords across technological codes in machinery. We look at the share of automation keywords ("all" in the table) and then focus on the three main subcategories, namely automat* patents, robot patents and numerical control (CNC) patents (defined above). The 95^{th} and 90^{th} percentile for the share of automation patents in the distribution of 6 digit codes in machinery define the threshold used to categorize auto95 and auto90 patents. The distributions are quite similar for the C/IPC 6 digit codes and for pairs of IPC 4 digit codes (see also the histograms below). As expected, the distributions are significantly shifted to the right for combinations of IPC 4 digit codes with G05 or G06. The distributions of each subcategory are right-skewed particularly for 6 digit codes and 4 digit pairs, and even more for the robot and CNC patents. The automat* keywords are also more common as the mean share for automat* is significantly higher than for the other keywords, however the difference narrows somewhat in the right tail: the 95^{th} percentile for 6 digit codes is 29.4% for the share of automat* patents and 13.7% and 12.7% for the share of robot and CNC patents. In the right tail, the distribution of robot patents and CNC patents are quite similar.

Figure 1.8.10 gives the histograms of the prevalence of automation keywords for all pairs of C/IPC 4 digit codes (panel a) and all pairs with at least one member in the machinery technological field (panel b). The histograms are very similar to those of C/IPC 6 digit codes in Figure 1.1. Figure 1.8.11 shows the histograms for all combinations of IPC 4 digit codes with G05 or G06 (panel a), or when the IPC 4 code is in the relevant technological field (panel b). Both distributions are considerably shifted to the right, in line with expectations since G05 proxies for control and G06 for algorithmic, two set of technologies which have been used heavily in automation. There are, however, much fewer combination of these types (in part because all histograms only consider groups with at least 100 patents), and accordingly few patents can be characterized as automation innovations this way.

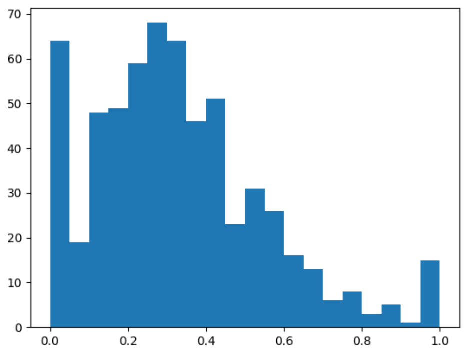


(a) For all pairs of C/IPC 4 digit codes

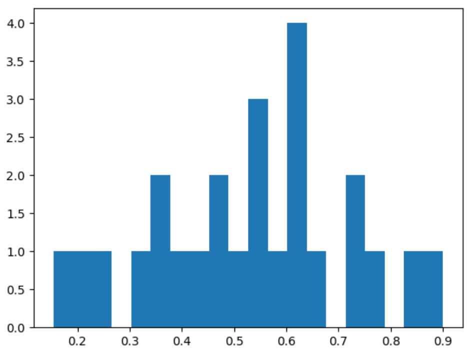


(b) For all pairs of C/IPC 4 digit codes within machinery with 100 patents

Figure 1.8.10: Histogram of the prevalence of automation keywords for C/IPC pairs of 4 digit codes



(a) For all combinations of IPC4 with G05 G06



(b) For combinations of IPC4 in machinery with G05 G06 and at least 100 patents

Figure 1.8.11: Histogram of the prevalence of automation keywords for combinations of IPC 4 digit codes with G05 G06

Table 1.8.35: Identification of automation technological categories

(a) Type of C/IPC codes identifying auto90 and auto95 patents

Ipc codes / Patents	Auto90	Auto95
Matches ipc6	78.2%	78.7%
Matches ipc4 pair	17.3%	24.3%
Matches ipc4 - G05/G06 combination	47.7%	47.8%

Note: Share of innovations classified as automation innovation through ipc6 codes, ipc4 pairs or ipc4 - G05/G06 pairs. Statistics computed on biadic patents from 1997-2011.

(b) Auto patents and subcategories of automation innovations

Sources / Patents	Auto80	Auto90	Auto95
Auto80	100.0%	100.0%	100.0%
Automat*80	36.2%	53.1%	72.1%
CNC80	5.0%	8.0%	13.2%
Robot80	12.0%	19.2%	33.6%
Auto90	62.4%	100.0%	100.0%
Automat*90	21.6%	34.6%	56.0%
CNC90	2.2%	3.6%	6.3%
Robot90	7.8%	12.5%	21.8%
Auto95	35.8%	57.3%	100.0%
Automat*95	4.4%	7.1%	12.4%
CNC95	1.6%	2.5%	4.4%
Robot95	6.3%	10.2%	17.7%

Note: Share of auto95 (auto90 and auto80, respectively) innovations which are also classified as automat*80/90/95, CNC80/90/95, and robot80/90/95 innovations. Statistics computed on biadic patents from 1997-2011.

How are auto90 and auto95 patents identified?

Given that our classification procedure is relatively complex, we assess here which features dominate. To do so, we focus on the set of $15,212,134$ biadic patent applications in 1997-2011 (corresponding to the $3,187,536$ patent families which have patent applications in at least two countries), since this corresponds to the set on which we run our main regressions. There are $310,458$ auto95 patent applications corresponding to $61,788$ patent families (and similarly $541,693$ auto90 patent applications corresponding to $107,237$ patent families). Table 1.8.35.a gives the share of biadic patents which are identified through a C/IPC 6 digit code, a pair of 4 digit codes or a combination of 4 digit code with G05/G06 (the shares sum up to more than 100% since patents may be identified as automation innovations in several ways). 6 digit codes appear to be the most relevant since they are enough to identify close to 80% of auto90 or auto95 patents alone.

Similarly, one may wonder which keywords are the most important in identifying automation patents. To do that, we define robot95 (respectively CNC95 or autm95) patents as patents which contain a technological group with a share of "robot" (respectively CNC or automat*) keywords above the threshold used to define auto95 (namely 0.4766), therefore those patents are a subset of the auto95 patents. We

Table 1.8.36: Confusion table for different classification periods

Confusion Matrix		Auto95 based on the 1998-1997 classification		Auto95 based on the 1998-2017 classification		Auto95 based on the 1997-2011 classification		Total
		Yes	No	Yes	No	Yes	No	
Auto95 based on	Yes	240,194	70,264	280,047	30,411	262,972	47,486	310,458
the 1978-2017	No	53,137	14,848,539	25,186	14,876,490	26,368	14,875,308	14,901,676
classification	Total	293,331	14,918,803	305,233	14,906,901	289,340	14,922,794	15,212,134

Notes: The statistics are always computed on patents from 1997-2011.

define robot90, CNC90, autm90, robot80, CNC80 and autm80 similarly. The other keywords are much less common. Table 1.8.35.b reports the share of auto95, auto90 and auto80 patents which belong to each subcategory. "Automat*" appears to be the most important keywords since 72% of auto95 patents are also automat*80 patents. "Robot" matters as well with 33.6% of auto95 patents which are robot80. This is true particularly at the top of the distribution since 17.7% of auto95 patents are also robot95 (more than autm95). CNC does not matter too much: only 13% of auto95 patents are CNC80.

Stability of the classification

To assess the stability of our classification, we redo exactly the same exercise but instead of using EPO patents from 1978 to 2017, we restrict attention to EPO patents from the first half of the sample (1978-1997), the second half of the sample (1998-2017) and the period of our main regression analysis (1997-2011). We focus on the same set of biadic patent applications in 1997-2011. Table 1.8.36 shows confusion tables on the classification of patents as auto95 according to each of the classification period. Regardless of the time period used the number of automation patents stays roughly constant. In particular, 85% of the baseline auto95 patents are still auto95 if we run the classification over the years 1997-2011. This common set of patents then represent 91% of all biadic patents classified as auto95 patents when using the period 1997-2011 instead of the full sample.

1.8.3 Redoing ALM

In this Appendix, we provide details on the analysis conducted in section 1.2.6. We use granted patents at the USPTO between 1970 and 1998. To assign patents to sectors, we first use Lybbert and Zolas (2014) who provide a concordance table between IPC codes at the 4 digit level and NAICS 1997 6 digits industry codes (mostly in manufacturing). The concordance table is probabilistic (so that each code is associated with a sector with a certain probability). In this exercise we are interested in matching patents with a sector of use and not the inventing sector (which is what is provided by the Eurostat concordance table for instance). The Lybbert and Zolas concordance tables are derived by matching patents texts with industry descriptions, and as such they cannot *a priori* distinguish between sector of use and industry of manufacturing. We checked, however, that patents associated with "textile and paper machines" for instance are associated with the textile and paper sectors and not with the equipment sector (as is the case with the Eurostat concordance table). We attribute patents to sectors fractionally in function of their IPC codes. To assign patents to the consistent Census industry codes used by ALM, we first use a Census concordance table (https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html) to go from NAICS 1997 to Census industry codes 1990, then we use the concordance table of ALM to get to the consistent Census industry codes of ALM. Finally, for each sector and each time period, we compute the sums of automation patents and machinery patents and take the ratio to be our measure of automation intensity. We exclude sectors with less than 50 machinery patents (which is why the number of sectors varies across time periods). We are left with 66 to 68 sectors, with only 7 of them not in manufacturing.

The other variables are directly taken from ALM. We refer the reader to that paper for a detailed explanation. The task measures are computed using the 1977 *Dictionary of Occupational Titles* (DOT) which measure the tasks content of occupations. Occupations are then matched to industries using the Census Integrated Public Micro Samples one percent extracts for 1960, 1970 and 1980 (IPUMS) and the CPS Merged Outgoing Rotation Group files for 1980, 1990 and 1998 (MORG). The task change measure at the industry level reflects changes in occupations holding the task content of each occupation constant, which ALM refer to as the extensive margin. Since tasks measures do not have a natural scale, ALM converted them into percentile values corresponding to their rank in the 1960 distribution of tasks across sectors, so that the employment-weighted means of all tasks measure across sectors in 1960 is 50. Our analysis only uses manufacturing sectors and starts in 1970 but we kept the original ALM measure to facilitate comparison. As in ALM, the dependent variable in Table 1.4 corresponds to 10 times the annualized change in industry's tasks inputs to favor comparison across periods of different lengths. Computerization ΔC_j is measured as the annual change in the percentage of industry workers using a computer at their jobs between 1984 and 1997 (estimated from the October Current Population Survey supplements), multiplied by 10 to ensure that all variables are over the same time length. For all regressions, observations are weighted by the employment share in each sector. In Table 1.4, the ratio of high-skill to low-skill workers are measured as the ratio of college graduates (and more than college) to high-school dropouts and graduates, taken from ALM—knowing that their data in turn come from IPUMS and MORG.

Table 1.8.38 reproduces Table 1.4 but with the laxer auto90 measure. The results are very similar—the only difference is that the coefficient on routine manual tasks is not significant at the usual levels in the 90s.[42]

Table 1.8.39 reproduces the Table 5 of ALM by carrying the analysis of Table 1.4 for each education groups over the time period 1980-1998 with the auto95 measure (the results are very similar with auto90). The table shows that automation reduces the amount of routine tasks undertaken by high-school dropouts and high-school graduates. Following ALM, Panel F computes the average effect of automation in tasks changes (from Panel A) and how much of this average effect can be explained by changes within educational groups (from Panels B to E). We find that changes within educational categories explain a significant share of the overall reduction in routine tasks but changes in educational composition also play a role, in line with Column 6 of Table 1.4. In contrast, ALM found that nearly all of the decline in routine tasks due to computerization came from within educational group changes.

[42] To interpret the effect of the automation variable, note that the means are 0.13, 0.15 and 0.14 in the 70s, 80s and 90s, and the standard deviations are 0.10, 0.12 and 0.11 with the auto90 definition.

ind6090	Title	Auto95	ind6090	Title	Auto95
16	Ag production crops and livestock; Ag services; Horticultural services	0.026	211	Other rubber products and plastics footwear and belting + tires and inner tubes	0.010
30	Forestry	0.035	212	Misc. plastic products	0.019
31	Fishing, hunting and trapping	0.013	220	Leather tanning and finishing	0.014
40	Metal mining	0.023	221	Footwear, except rubber and plastic	0.086
41	Coal mining	0.037	222	Leather products, except footwear	0.014
42	Crude petroleum and natural gas extraction	0.021	230	Logging	0.030
50	Nonmetallic mining and quarrying, except fuel	0.048	231	Sawmills, planning mills, and millwork	0.038
66	Construction	0.036	236	Railroad locomotives and equipment; Cycles and misc transporation equipment; Wood buildings and mobile homes	0.109
100	Meat products	0.107			
101	Dairy products	0.402			
102	Canned and preserved fuits and vegetables	0.007	241	Misc. wood products	0.075
110	Gain mill products	0.030	242	Furniture and fixtures	0.043
111	Bakery products	0.005	246	Scientific and controlling instruments; Optical and health service supplies	0.410
112	Sugar and confectionary products	0.022			
120	Beverage industries	0.017	250	Glass products	0.017
121	Misc. food preparations, kindred products	0.019	251	Cement, concrete, gypsum and plaster products	0.074
130	Tobacco manufactures	0.033	252	Structural clay products	0.033
132	Knitting mills	0.007	261	Pottery and related products	0.027
140	Dyeing and finishing textiles, except wool and knit goods	0.004	262	Misc. nonmetallic mineral and stone products	0.038
			270	Blast furnaces, steelworks, rolling and finishing mills	0.039
141	Floor coverings, except hard surfaces	0.009			
142	Yarn, thread, and fabric mills	0.071	271	Iron and stell foundaries	0.178
146	Primary aluminum & other primary metal industries	0.083	281	Cutlery, handtools, and other hardware	0.023
			282	Fabricated structural metal products	0.034
150	Misc. textile mill products	0.079	346	Plastics, synthetics and resins; Soaps and cosmetics; Agricultural chemicals; Industrial and miscellaneous chemicals	0.028
151	Apparel and accessories, except knit	0.060			
152	Misc. fabricated textile products	0.172			
160	Pulp, paper, and paperboard mills	0.020	351	Transportation equipment	0.207
161	Misc. paper and pulp products	0.015	360	Ship and boat building and repairing	0.058
162	Paperboard containers and boxes	0.003	362	Guided missiles, space vehicles, ordnance, aircraft and parts	0.166
166	Screw machine products; Metal forgings & stampings; Misc. fabricated metal products	0.086			
			380	Photographic equipment and supplies	0.043
172	Printing, publishing, and allied industries except newspapers	0.017	381	Watches, clocks and clockwork operated devices	0.174
176	Engine and turbines; Construction and material handling machines; Metalworking machinery; Machinery, except electrical, n.e.c.; Not specified machinery	0.125	391	Misc. manufacturing industries and toys, amusement and sporting goods	0.032
			460	Electric light and power	0.161
181	Drugs	0.040	462	Eletric and gas, and other combinations	0.153
186	Electronic computing equipment; Office and accounting machines	0.320	470	Water supply and irrigation	0.126
			471	Sanitary services	0.018
190	Paints, varnishes, and related products	0.015	636	Grocery stores; Retail bakeries; Food stores, n.e.c.	0.004
200	Petroleum refining	0.031			
201	Misc. petroleum and coal products	0.010			
206	Household appliances; Radio, TV &	0.221			

Table 1.8.37: List of sectors in the ALM regressions

Table 1.8.38: Changes in task intensity and skill ratio across sectors and automation (auto90)

	(1) Δ Nonroutine analytic	(2) Δ Nonroutine interactive	(3) Δ Routine cognitive	(4) Δ Routine manual	(5) Δ Nonroutine manual	(6) Δ H/L
Panel A: 1970 - 80, n=67						
Share of automation patents in machinery	0.82	3.57	-17.95***	-10.60***	-0.89	0.11**
	(3.51)	(4.32)	(4.22)	(3.74)	(5.13)	(0.05)
Δ Computer use 1984 - 1997	-7.16	-2.99	-18.91***	-3.26	14.86*	0.08
	(5.71)	(7.03)	(6.86)	(6.09)	(8.36)	(0.09)
Intercept	0.92	2.14*	4.34***	3.39***	-1.70	0.04***
	(1.00)	(1.23)	(1.20)	(1.07)	(1.47)	(0.02)
R^2	0.02	0.01	0.31	0.12	0.05	0.08
Weighted mean Δ	-0.05	2.17	-0.90	1.49	0.42	0.07
Panel B: 1980 - 90, n=67						
Share of automation patents in machinery	9.01*	13.29**	-25.37***	-13.79***	9.70**	0.73***
	(5.41)	(6.23)	(4.96)	(4.28)	(4.72)	(0.19)
Δ Computer use 1984 - 1997	24.75**	22.95*	-13.41	-1.55	-5.37	0.39
	(10.34)	(11.90)	(9.49)	(8.18)	(9.02)	(0.37)
Intercept	-3.15*	-1.21	3.55**	1.69	-2.39	-0.06
	(1.77)	(2.03)	(1.62)	(1.40)	(1.54)	(0.06)
R^2	0.13	0.13	0.32	0.14	0.06	0.21
Weighted mean Δ	1.86	4.17	-2.22	-0.59	-1.74	0.11
Panel C: 1990 - 98, n=67						
Share of automation patents in machinery	9.23**	10.63*	-13.47***	-6.24	3.95	0.42***
	(4.57)	(6.22)	(5.12)	(4.19)	(4.76)	(0.12)
Δ Computer use 1984 - 1997	27.31***	28.19**	-25.09***	-26.11***	8.05	0.73***
	(8.27)	(11.25)	(9.26)	(7.58)	(8.61)	(0.22)
Intercept	-2.93**	-1.93	2.23	2.41*	-2.55*	-0.08**
	(1.44)	(1.96)	(1.61)	(1.32)	(1.50)	(0.04)
R^2	0.20	0.14	0.20	0.19	0.03	0.29
Weighted mean Δ	2.45	3.79	-3.44	-2.36	-0.79	0.09

Standard errors are in parentheses. Colums (1) to (5) of Panels A to C each presents a separate OLS regression of ten times the annual change in industry-level task input between the endpoints of the indicated time interval (measured in centiles of the 1960 task distribution) on the share of automation patents in machinery (defined with the 90th percentile threshold) and the annual percentage point change in industry computer use during 1984 - 1997 as well as a constant. In Column (6), the dependent variable is the ratio of high-skill (college graduates) to low-skill (high-school graduates and dropouts) workers. Estimates are weighted by mean industry share of total employment in FTEs over the endpoints of the years used to form the dependent variable. * p<0.1; ** p<0.05; *** p<0.01

Table 1.8.39: Changes in task intensity and skill ratio across sectors and automation (auto95) by skill groups

	(1) Δ Nonroutine analytic	(2) Δ Nonroutine interactive	(3) Δ Routine cognitive	(4) Δ Routine manual	(5) Δ Nonroutine manual
Panel A: Aggregated within-industry change					
Share of automation patents in machinery	9.53**	17.97***	-26.66***	-17.09***	12.57***
	(4.53)	(5.39)	(4.83)	(3.90)	(4.30)
Δ Computer use 1984 - 1997	24.91***	23.81***	-17.75***	-11.53**	0.47
	(6.36)	(7.56)	(6.79)	(5.48)	(6.03)
Intercept	-2.36**	-1.01	2.05*	1.73*	-2.37**
	(1.03)	(1.22)	(1.10)	(0.89)	(0.98)
R^2	0.26	0.27	0.39	0.29	0.12
Weighted mean Δ	2.05	3.88	-2.62	-1.29	-1.34
Panel B: Within industry: High school dropouts					
Share of automation patents in machinery	2.41	13.61	-26.19***	-5.80	4.56
	(7.89)	(10.85)	(6.94)	(6.22)	(6.35)
Δ Computer use 1984 - 1997	11.70	18.08	15.84	8.68	-9.95
	(11.08)	(15.24)	(9.74)	(8.73)	(8.91)
Intercept	-4.47**	-8.45***	0.87	0.55	1.16
	(1.79)	(2.47)	(1.58)	(1.41)	(1.44)
R^2	0.02	0.05	0.19	0.02	0.02
Weighted mean Δ	-2.56	-4.73	1.20	1.39	0.04
Panel C: Within industry: High school graduates					
Share of automation patents in machinery	-7.08	6.50	-26.09***	-13.43***	9.62*
	(5.47)	(7.05)	(5.64)	(4.25)	(5.37)
Δ Computer use 1984 - 1997	9.30	-0.76	-14.39*	-2.86	6.71
	(7.69)	(9.90)	(7.92)	(5.96)	(7.54)
Intercept	-2.86**	2.19	2.25*	0.00	-1.43
	(1.24)	(1.60)	(1.28)	(0.97)	(1.22)
R^2	0.04	0.01	0.30	0.14	0.06
Weighted mean Δ	-2.03	2.57	-1.88	-1.45	0.30
Panel D: Within industry: Some College					
Share of automation patents in machinery	-11.94	-7.49	-4.92	-5.92	12.48*
	(8.04)	(7.31)	(6.01)	(5.72)	(6.56)
Δ Computer use 1984 - 1997	7.05	13.85	-14.68*	-14.11*	9.14
	(11.29)	(10.26)	(8.44)	(8.03)	(9.20)
Intercept	-1.10	0.31	0.38	2.21*	-2.74*
	(1.83)	(1.66)	(1.37)	(1.30)	(1.49)
R^2	0.04	0.04	0.06	0.07	0.07
Weighted mean Δ	-0.97	1.78	-2.17	-0.33	-0.43
Panel E: Within industry: College graduates					
Share of automation patents in machinery	-6.54	-7.28**	-11.58*	-7.70	17.00***
	(4.25)	(3.59)	(6.48)	(7.74)	(6.03)
Δ Computer use 1984 - 1997	14.44**	9.29*	-5.55	-7.69	11.14
	(6.00)	(5.06)	(9.14)	(10.91)	(8.50)
Intercept	-0.94	0.17	-1.22	-0.14	-5.35***
	(0.97)	(0.82)	(1.48)	(1.77)	(1.38)
R^2	0.01	0.09	0.06	0.03	0.14
Weighted mean Δ	0.69	0.99	-2.93	-1.86	-2.40
Panel F: Decomposition of automation effects into within and between education group					
Explained task Δ	0.73	1.38	-2.04	-1.31	0.96
Within educ groups (%)	-63.96	15.80	72.32	54.61	81.96
Between educ groups (%)	163.96	84.20	27.68	45.39	18.04

n in Panels A-D is 69 and in Panel E it is 68 consitent CIC industries. Standard errors are in parentheses. Each column of panels A - E presents a separate OLS regression of ten times the annual change in industry-level task input for the relevant education group (measured in centiles of the 1960 task distribution) during 1980 - 1998 on the the share of automation patents in machinery (defined with the 95th percentile threshold) and the annual percentage point change in industry computer use during 1984 - 1997 as well as a constant. Estimates are weighted by mean industry share of total employment (in FTEs) in 1980 and 1998. The 'explained' component in Panel F is the within-industry change in the task measure predicted by the share of automation patents in regression models in Panel A. * p<0.1; ** p<0.05; *** p<0.01

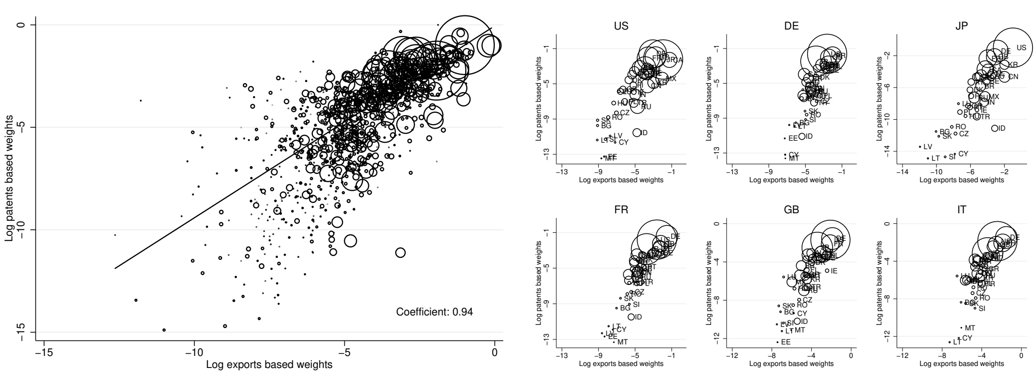
1.8.4 Validating our weights approach

In this Appendix, we compare our firm-level weights to bilateral trade flows and show that they are strongly correlated. The first step is to compute patent-based weights at the country level. For this exercise (and this exercise only), we define the domestic country d of a firm based on the location of its headquarters (according to the country code of its identifier in the Orbis database—for firms which we merged, we keep the country code of the largest entity by biadic machinery patents in 1997-2011). We compute the foreign weights for each firm i by excluding the domestic country. Therefore the foreign weight for country $c \neq d$ for firm i is given by $\omega_{i,c}/(1-\omega_{i,d})$ (recall that these weights are computed based on patenting from 1970 to 1994). We then build the foreign patent-based weight in country c for country $d$as a weighted average of the foreign weights in country c of the firms from country d (each firm is weighted according to the number of machinery biadic patents in 1997-2011).

The second step is to build similar weights based on exports. To do that, we collect sectoral bilateral trade flow from UN Comtrade data between between 1995 and 2009 for 40 countries (Taiwan is not included in the data). To obtain trade flows in machinery, we use a concordance table between 4 digit IPC codes and 2 or 3 digits NACE Rev 2 codes provided by Eurostat, this concordance table matches IPC codes to the industry of manufacturing. The concordance table assigns a unique industry to each IPC code. Then, for each industry and each country, we compute the share of patents over the period 1995-2009 which are in machinery according to our definition.[43] This gives us a machinery weight for each industry code and each country. We then multiply sectoral trade flows (after having aggregated the original data to the NACE Rev 2 codes used in the concordance table) by this weight to get bilateral trade in machinery. We then compute the export share in machinery across destinations. We could compute trade based weights for each year but here we report results based on 1996 only (there are a few missing observations for 1995).

Figure 1.8.12 plots the patent-based weights against the trade-based weights.

[43]To do that we use a fractional approach: each patent is allocated NACE sectoral weights (and machinery weights) depending on the share of IPC codes associated with a NACE sector or machinery.



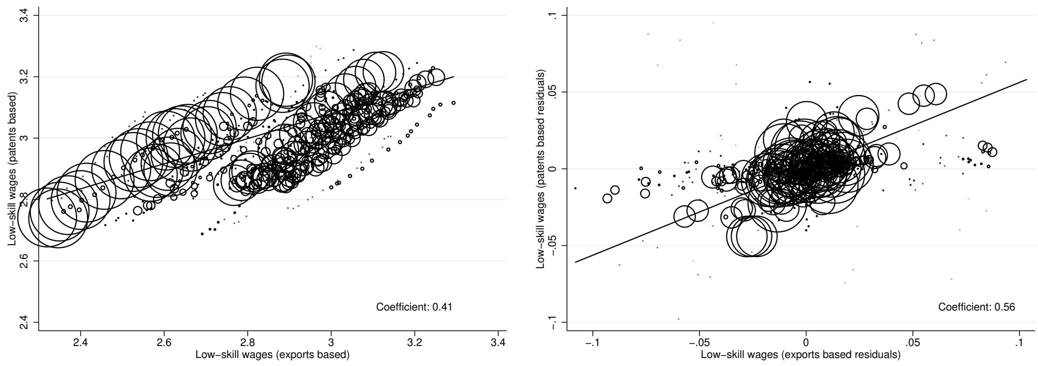
(a) All countries **(b)** Trade from the 6 largest countries

Figure 1.8.12: Bilateral patent flows and trade flows in machinery. Panel (a) plots log patent based weights, which are a weighted average of the destination country's weights in the (foreign) patent portfolio of firms from the origin country, against export shares in machinery over the years 1995-2009. The size of each circle represents the product of the GDP of both countries, which is used as a weight in the regression. Panel (b) focuses on the weights from the listed countries and observations are weighted by the GDP of the partner country.

Panel (b) focuses on a few origin countries while Panel (a) plots all countries together. We find a strong correlation between the two measures with a regression coefficient of 0.94 (when observations are weighted by the trade flow in 1996).

Another way to summarize how close the two distributions are is to compute what low-skill wages would be according to either sets of weights. We do this in Figure 1.8.13. There for each country, we compute "foreign low-skill wages" as a weighted average of foreign wages where the weights are either the patent-based weights or the trade-based weights derived above. Foreign wages are deflated with the local PPI and converted in USD in 1995 as in our main analysis. Panel (a) then reports foreign log low-skill wages according to both types of weights in 1995-2009, we find that they are strongly correlated. Panel (b), reports the same foreign log low-skill wages but taking away country and year fixed effects. We find a regression coefficient of 0.56, when observations are weighted by the number of machinery patent in the country over the 1997-2011 time period.

Overall, this exercise shows that there is tight relationship between our patent-based weights and (future) trade flows, suggesting that we can use these patent-based weights as proxies for firms' markets exposure.



(a) Low-skill wages. **(b)** Residualized low-skill wages

Figure 1.8.13: Foreign low-skill wages for each country computed either with patent-based weights or with trade-based weights. Wages are computed for the years 1995-2009. Panel (a) plots log foreign low-skill wages using either patent-based weights or trade-based weights. Panel (b) plots the residuals of foreign wages according to both methods controlling for country and year fixed effects. Observations are weighted by the number of biadic machinery patents by firms from the the country over the years 1997-2011.

1.8.5 Macroeconomic variables

Our main source of macroeconomic variables is the *World Input Output Database (WIOD)* from Timmer, Dietzenbacher, Los, Stehrer and de Cries (2015) which contains information on hourly wages (low-skill, middle-skill and high-skill) for the manufacturing sector and the total economy from 1995 to 2009 for 40 countries. It further contains information on both GDP deflators and producer price indices both for manufacturing and for the whole economy. Their data on skill is based on the 1997 International Standard Classification of Education (ISCED) system, where category 1+2 denote low-skill (no high-school diploma in the US) 3+4 denote middle-skill (high-school but not completed college) and 5+6 denotes high-skill (college and above). Switzerland is not included in the WIOD database and we add data on skill-dependent wages, productivity growth and price deflators manually using data obtained directly from *Federal Statistical Office of Switzerland.*

We supplement this data with data from *UNSTAT* on exchange rates and GDP (and add Taiwan separately from the *Taiwanese Statistical office*). We use this data to calculate the GDP gap as the deviations of log GDP from HP-filtered log GDP using a smoothing parameter of 6.25.

The primary data source for the hourly minimum wage data is *OECD Statistics.* Not all countries have government-imposed hourly minimum wages. Spain, for instance, had a monthly minimum wage of 728 euros in 2009. To convert this into hourly wage we note that Spain has 14 monthly payments a year (+1 payments in December and July). Further, workers have 6 weeks off and the standard work week is 38 hours. Consequently we calculate the hourly minimum wages as monthly minimum wage$\times 14/\left[(52-6)\times 38\right]$, which in the case of 2009 is 5.83 euros per hour. We perform similar calculations, depending on individual work conditions, for other countries with minimum wages that are not stated per hour: Belgium, Brazil, Israel, Mexico, Netherlands, Poland and Portugal.

For the US, we use data from FRED for state minimum wages and calculate the nation-level minimum wage as the weighted average of the state-by-state maximum of state minimum and federal minimum wages, where the weight is the manufacturing employment in a given state.

Further, the UK did not have an official minimum wage until 1999. Correspondingly, we follow Dickens, Machin and Manning (1999) and use the wage levels agreed upon by local wage councils. These were in effect from 1909 until 1993. For, 1995-1998, the four years in our sample where no official minimum wage existed, we use the nominal level from 1993. We use the employment-weighted industry average across manufacturing industries. Finally, Germany did not have a minimum wage during the time period we study. Instead, we follow Dolado, Kramarz, Machin, Manning, Margolis and Teulings (1996) and use the collectively bargained minimum wages in manufacturing which effectively constitute law once they have been implemented. These data come from personal correspondence with the Sabine Lenz at the *Statistical Agency of Germany.*

Chapter 2

Automation and the Labor Share: Evidence from Patents

This chapter is joint work with Vladimir Sulaja.*

2.1 Introduction

Since the 1980s, the labor share of aggregate income has been decreasing steadily over the majority of countries. Karabarbounis and Neiman (2014) find that the corporate labor share has decreased by around five percentage points over the last 35 years. This decrease in the labor share has important implications for inequality. Most people make a living from their labor, and if capital is concentrated in the hands of a few, a reduction in the share of income going to labor means that the majority of people are relatively worse off.

Despite its importance, we are still far away from understanding the causes of the decline we have witnessed. An explanation that is popular in the public discourse is the rising adoption of automation technologies. Automation technologies include industrial robots and artificial intelligence but also a wide range of computer and software technology which makes manual work redundant.

*Vladimir Sulaja: University of Zurich, Ramistrasse 59, vladimir.sulaja@econ.uzh.ch. Carlo Zanella: University of Zurich, Ramistrasse 59, carlo.zanella@econ.uzh.ch. We thank David Hémous, Lei Li, Tong Zhang, as well as various participants at the macro seminar at the University of Zurich for helpful comments and suggestions.

The idea that automation could have enormous consequences for the labor market dates back to at least Keynes (1930), who argued that the advance of technology would lead to technological unemployment. The logic is simple — as the technology advances, robots will take over production and many will be left jobless. Of course, this prediction has not taken into account that many jobs would be created by new technologies. While the scenario that Keynes (1930) imagined has not realized, the world has seen a substantial change in the labor markets due to technology.

In this paper, we study the relationship between the fall in the labor share over the last three decades with the onset of automation technologies. We use the patent classification developed by Dechezleprêtre et al. (2019) to conduct our empirical analysis. Our measure of automation innovation is the number of automation patents in a country in a year. While there has of course been previous research on the relationship between automation and the labor share, we are the first to use automation patents as a measure for automation technology. Autor and Salomons (2018), for instance, use total factor productivity as a measure of automation, whereas Acemoglu and Restrepo (2017b) focus on robots. Our patent-based measure has the advantage of being a direct measure of automation. Moreover, it is more comprehensive than robots.

While the measure of automation is more general than in previous studies, the patent classification by Dechezleprêtre et al. (2019) has been explicitly developed for the machinery sector. The definition of automation in this paper therefore does not include computer technologies outside the use in manufacturing. It would, for instance, not capture advances in Artificial Intelligence (AI). However, A.I. innovation only started to take off after 2012 (WIPO, 2019), whereas the decline in the labor share has been going on since the 1980.

In the main part of the paper, we document a robust negative association between automation and the labor share. Our main results suggest that an increase in the number of automation patents by ten percent is associated with a decrease in labor share by 0.16–0.31 percentage points. For the U.S., for instance, this would imply that automation accounts for 48 to 91 percent of the decline of the corporate labor share.

We use the data from World Input Output Database (WIOD) to separate the labor share of income in three categories: low-skilled, middle-skilled, and high-skilled labor. This data also features the possibility of dividing the labor share of each skill group into hours worked and wages.

We show that the relationship is entirely driven by the middle-skilled labor share, which is consistent with Autor and Dorn (2013b) and related literature who document shifts from middle-skilled to low- and high-skilled employment.

There is no significant correlation between low- and the high-skilled labor share with automation. We find that automation is negatively correlated with overall employment. Moreover, it is associated with a decrease in the share of hours of low-skilled labor and an increase in the share of high-skilled workers. Surprisingly, there is no significant association with the share of hours worked by middle-skilled labor. This suggests that automation displaces low-skilled workers but is complementary to high-skilled labor. This view is consistent with the theory put forth by previous research such as Hémous and Olsen (2018).

Our main focus is on establishing the sign and the size of the statistical relationship between automation and the labor share. We are careful about interpreting our results as the causal effect of automation on the labor share. It is likely that automation innovation and production are jointly determined. For instance, Dechezleprêtre et al. (2019) show that higher low-skilled wages lead firms to develop more labor-saving innovation. Zator (2019) studies the effect of automation on worker substitution by robots using data on German firms and finds that endogeneity leads to a significant bias in the coefficient on the effect of automation on worker displacement.

For our empirical analysis, we use data on the corporate labor share (CLS) collected by Karabarbounis and Neiman (2014). Calculating the labor share in the entire economy requires splitting business income between labor and capital, which is not straightforward and requires making additional assumptions. In the corporate sector this is not an issue, which is why we focus on the corporate labor share.

Our analysis focuses on OECD countries in the time period 1995–2012 because for periods before 1995 and after 2012 the data on the corporate labor share is spotty. We show, however, that the negative relationship between automation and the labor share holds in larger samples and for various measures of the labor share.

Our work is also related to the recent literature on the job polarization, which refers to the shrinking share of employment in middle-skilled, routine occupations. Autor et al. (2003b) argue that the job polarization is a consequence of the progress of automation. Cortes et al. (2017) use a neoclassical framework to study this problem and find that the automation accounts for a very small share of job polarization. We find that while automation is not associated with a decrease in middle-skilled employment, there is a negative association between middle-skilled labor share and automation. Graetz and Michaels (2017) study the relationship between modern technology and jobless recoveries, which Cortes et al. (2017) argue are tightly related to job polarization. Using country-level data he finds that industries that used more routine tasks and higher exposure to robotization did not experience slower employment recoveries. Graetz and Michaels (2018b) study robots adoption and find that it increases total factor productivity. They show that although increased use of robots does not significantly reduce total employment, it reduces the low-skilled employment share. Martinez (2019) uses industry-level data and a new measure of aggregate task inputs in production and finds evidence that automation was a significant driving force of the US labor share between 1972–2010. Our work is different as we study automation and labor share on the country level and use patents as a measure of automation rather than task inputs.

Of course, we are not the first to study the causes of the decline of the labor share. Karabarbounis and Neiman (2014) investigate the possibility that the decreasing relative price of investment with respect to consumption is responsible for the decline in the labor share. In their model, firms produce two types of goods using a continuum of intermediate inputs. Monopolistically competitive firms produce intermediate inputs using capital and labor. Shifts in technology differences make the rental rate of capital lower, which forces firms to switch labor for capital and decrease the share of labor in GDP.

Kehrig and Vincent (2018) use micro-level data from U.S. manufacturing to investigate the properties of the decline we witnessed. They find that since the mid-1980s, the labor share in U.S. manufacturing declined around five percentage points. However, the median establishment has increased its labor share by around 1.4 percentage points per decade. Looking further into the forces at play, they find that the hyperproductive plants have over time captured a higher share of the industry, and that these firms are also the ones that have a lower labor share.

Similarly, Autor et al. (2019) study the fall of labor share in the context of the rise of superstar firms using industry-level data. They find that superstar firms that increase their market shares are also the ones that have a lower labor share. This lead to an overall decline in labor share.

Elsby et al. (2013) argue that the main cause of the decreasing labor share was offshoring activity. They argue that explanations for the decline in the labor share that rely on the aggregate perspective are not supported by the data. For example, the decrease in the relative price of investment could not be the explanation for the decreasing labor share because in times of large shifts in the labor share, rental rate of capital does not increase significantly.

Eden and Gaggl (2018) study the effects of information and communication technologies on payments to factors of production. They document that the income share of ICT in GDP has increased sevenfold, while the share of other types of capital has remained stable. Additionally, they find that the reallocation of income from labor to capital happened in occupations which are highly substitutable by ICT.

Koh et al. (2018) investigate the role of intellectual property products in the decline of the labor share. Unlike most studies, they use data on the labor share spanning 65 years. They show that the decline in the labor share can be explained entirely by the increasing importance of intellectual property products in national income. However, their study only includes the data on the US.

Rognlie (2015) focuses on the share of income that accrues to capital, and finds that the increase in the net capital share since 1948 comes entirely from the housing sector. He goes on to argue that in order for the capital accumulation to explain the trends in capital share, we would need much different elasticities of substitution and correlation between the capital-income ratio and capital share that we do not observe.

Boehm et al. (2019) use the establishment-level data from multinational firms in the U.S. manufacturing sector. Multinational establishments accounted for 41 percent of aggregate employment decline. They find that newly multinational establishments in the U.S. experienced job losses, while their parent firms increased imports from their parent firms abroad.

Barkai, Simcha (2016) finds that markups have grown over time and that this has led to both a decrease in the labor and capital share.

The paper is organized as follows. Section 2 presents the data on labor share, decomposition of labor share and automation innovation, as well as descriptive statistics. Section 3 documents the effect of automation innovation on labor share both overall and by skill groups, and Section 4 concludes.

2.2 Data

In this section, we describe the data that we use for our analysis. We first present our measure of automation and then describe the sources of our macroeconomic variables. Finally, we present some summary statistics.

2.2.1 Automation measure

For our measure of automation, we use the patent classification developed by Dechezleprêtre et al. (2019), who use text analysis in order to categorize technological codes assigned to patents, called IPC and CPC codes, as automation.

We briefly summarize the key facts from the classification. The classification of patents relies on the fact that every patent is associated with technological codes, called IPC/CPC codes. Dechezleprêtre et al. (2019) employ a full-text keyword search on the text of EPO patents. For every technological code they compute the frequency of patents that match a certain set of keywords, such as “robot”, “automation”, “Computer Numerical Control”, etc. They then classify a patent from the universe of patents in PATSTAT as automation if it is assigned an IPC/CPC code with a frequency of matching patents that is higher than a given threshold.

Dechezleprêtre et al. (2019) introduce both a stricter and a laxer classification of automation patents depending on the threshold. We use both the 90th or the 95th percentile of the IPC/CPC code distribution as the threshold. The resulting classifications we call auto90 and auto95, respectively. Dechezleprêtre et al. (2019)

validate the measure of automation by including it in the exercise of Autor et al. (2003b) and find that in the U.S., sectors where the share of automation patents filed in machinery was high, are associated with a decrease in routine tasks and an increase in the skill ratio.

For our analysis, we use the PATSTAT Autumn 2018 database, which contains information on almost the entire universe of patents. We combine the automation classification with PATSTAT and consider automation patents which have been registered with at least two patent authorities. Following Dechezleprêtre et al. (2019), we call these patents "biadic".[1] Our focus is on biadic patents because previous research including, for instance, Henderson and Cockburn (1996) and, more recently, Dechezleprêtre et al. (2017), argues that biadic patents reflect more important innovations. However, we will show that our results are robust to alternative innovation quality indicators.

We combine the automation classification with PATSTAT and define as automation in a country and year the number of all biadic auto95 and auto90 patents which have been invented in that year and registered with the country's patent authority. We define the year of invention as the earliest application year within its patent family. EPO patents are assigned to countries upon entry into the national phase.

2.2.2 Macroeconomic data

For our empirical analysis, we use data on the labor share as well as other macroeconomic variables such as GDP, employment and labor compensation on a country-year level.

Data on the corporate labor share (CLS) are taken from Karabarbounis and Neiman (2014). We also use the labor share from the Penn World Tables as an alternative measure. We use real GDP from the World Bank. To control for business cycle fluctuations, we define the GDP gap as the residuals of applying the HP filter to real GDP. The World Input-Output Database (Timmer et al., 2015b) contains data on employment and total labor compensation by skill group. We define low-,

[1]Note that this definition is in line with DHOZ but slightly than in most of the literature which defines as biadic patents that were registered at at least two of the following authorities: EPO, USPTO, JPO.

middle-, and high-skilled wages as the ratio of the skill group's total labor compensation and hours worked by that skill group. Similarly, we define labor shares by skill group as the skill group's labor compensation divided by gross value added.

For our alternative measures of the labor share, we use data from EU KLEMS and from Penn World Table.

2.2.3 Sample

Our baseline analysis with the corporate labor share by Karabarbounis and Neiman (2014) includes 29 OECD countries and spans the years 1995-2012. We restrict the sample to OECD countries because most of them have a significant number of auto95 patents.[2] We will show in the appendix, however, that our results also hold in larger samples.

Table 2.5.1 shows summary statistics of our patent counts and the corporate labor share.

2.3 Empirical Analysis

In this section, we present the results of our empirical analysis. We will first present our panel regression results. Second, we present our long-difference results. Third, we show the results by different skill groups. The rest of the section demonstrates the robustness of our results.

2.3.1 Main results

In this section, we present our main empirical results. We focus on the corporate labor share because in the corporate sector the problem of distinguishing between an entrepreneur's capital income and labor income does not arise (Karabarbounis and Neiman, 2014). In our baseline regressions we estimate the following fixed-effects

[2]For instance Saudi Arabia, which is not in the OECD, has only between 0 to three auto95 patents in 1995–2012 despite having a relatively large GDP.

panel model:

$$CLS_{c,t+1} = \beta_0 + \beta_1 \ln \text{auto95}_{c,t} + X'_{c,t}\gamma + \delta_t + \delta_c,$$

where $CLS_{c,t+1}$ is the corporate labor share (multiplied by 100), $\ln \text{auto95}_{c,t}$ is the logarithm of the number of biadic auto95 innovations patented at country c in year t, $X_{c,t}$ is a vector of controls, and δ_t and δ_c are year and country fixed effects, respectively.

Table 2.3.1: Baseline results

	Corporate Labor Share				
	(1)	(2)	(3)	(4)	(5)
lnauto95	-2.292***	-1.646***	-3.124***	-2.739***	-2.702***
	(0.389)	(0.482)	(0.758)	(0.718)	(0.764)
lnother95			1.649*	1.303	1.037
			(0.902)	(0.832)	(2.073)
GDP gap				37.84***	37.89***
				(8.676)	(8.665)
lnnonmach					0.247
					(1.730)
Fixed effects	C	C+Y	C+Y	C+Y	C+Y
Observations	429	429	429	429	429
Countries	29	29	29	29	29

Note: Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Our baseline results are reported in Table 2.3.1. The dependent variable is the corporate labor share, and all the independent variables are lagged by one year. We explore different timing assumptions in Section 2.3.4. In Column (1), we run a simple univariate OLS regression without any controls except country fixed effects. The estimated coefficient suggests that a 10 percent increase in automation leads to a decrease in the labor share of 0.229 percentage points. In Column (2), we add year fixed effects and still find significant effects although the magnitude of the coefficient is somewhat lower. Next we address the issue that the negative relationship may be driven by a growth of innovations in the machinery sector in general rather than labor-displacing technologies. In Column (3), we control for other machinery innovations except auto95 (other95). Interestingly, we find a weakly significant positive influence of non-automation machinery innovation. The negative relationship between automation and the labor share, on the other hand, becomes more pronounced. It is well known that both the labor share and patenting is correlated with the business cycle (Hingley and Park, 2017; Botelho, 2018). To control for business cycle fluctuations, in Column (4), we add the GDP gap to the regression model. We find a large positive effect of the GDP gap on the labor share.[3] The coefficient on automation remains significant and remarkably stable. Finally, Column (5), controls for non-machinery innovation. The results suggest that non-machinery innovation is not significantly related to the labor share.

Overall, our results imply a significant relationship between automation and the corporate labor share. In the U.S., for instance, the number of automation patents registered per year has grown at an annualized rate of 8.5 percent between 1995 and 2011. Our results suggest that a decrease of 0.140–0.266 percentage points per year of the labor share can be explained by growth in automation innovation. This would account for 48 to 91 percent of the decline of the U.S. corporate labor share.

Note that we prefer auto95 as a measure of automation because it is stricter than auto90. The results on auto90 are similar and reported in Table 2.3.6 in the Appendix.

2.3.2 Long Differences

Both technological change and changes in the labor share reflect long-term processes. However, so far we identified the coefficient on automation in a yearly panel, which effectively exploits variation of a high frequency. In this section, we repeat our baseline analysis using low-frequency variation.

In Table 2.3.2, we regress five-year changes in the labor share on the logarithm of the number of automation patents. In Column (1), we report the results without any controls. The coefficient is negative and statistically significant. In Column (2), we add period fixed effects, and in Column (3) and (4), we add back non-automation machinery innovation and the GDP gap. In all specifications except (2), we confirm the significant negative estimates of the previous section. In specification

[3]The labor share has been considered countercyclical until recently. However, as Botelho (2018) shows, this has reversed in the last three decades.

(3) and (4), the magnitudes are even larger than in the yearly panel. This suggests that automation is negatively related to automation over periods spanning at least several years.

Table 2.3.2: Long Differences

	5-year Changes ΔCLS			
	(1)	(2)	(3)	(4)
Δlnauto95	-1.754***	-1.431	-3.655**	-4.026***
	(0.634)	(0.945)	(1.663)	(1.508)
Δlnother95			2.341	1.929
			(1.555)	(1.320)
ΔGDP gap				74.99**
				(30.48)
Fixed effects		Y	Y	Y
Observations	75	75	75	75
Countries	28	28	28	28

Note: The differences are constructed over the following time spans: 1995–2000,2000-2005, 2005-2011. Our last interval therefore spans 6 rather than 5 years. We choose 2011 as our end year because we lose a lot of countries in 2012. Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.3.3 Effects by Skill Groups

The routinization hypothesis put forward by Autor et al. (2003b) suggests that automation technologies displace workers in middle-skilled routine occupations — occupations which mostly involve tasks that can be done by following well-defined instructions. At the same time, automation technologies may be complementary to workers engaged in abstract tasks, which are found mostly in high-skilled occupations. The evidence on this hypothesis is mixed. For instance, Cortes et al. (2017) develop a model with endogenous participation and occupation choice and find that, quantitatively, automation can only explain a modest part of the decline in routine jobs. In this section, we study how automation is related to the labor share, employment, and wages of different skill groups.

In order to do that, we use data from the World Input-Output Database to construct the labor share by skill group. In the WIOD, skill is equivalent to educational attainment. We then regress these variables on our measure of automation. Our results are reported in Table 2.3.3. In Column (1), we regress labor share, wage, and employment, defined as the number of hours worked, on automation. Columns (2)–(4) report the results for low-, middle-, and high-skilled labor.

Our main result in this section is that the negative relationship of automation and the labor share is mainly driven by middle-skilled labor, as Panel A of Table 2.3.3 shows. This result is consistent with the routinization hypothesis. Surprisingly, we do not find a significant association of automation with the wages for any skill group. In Panel C. we show that automation associated with a decrease in total hours worked. Moreover, an increase in automation is associated with a decrease in the share of hours of the low-skilled and an increase in the share of hours of the high-skilled workers. There appears to be no effect on the relative share of work hours of the middle-skilled workers.

Table 2.3.3: Effects by Skill Groups

	(1) All	(2) Low-skilled	(3) Middle-skilled	(4) High-skilled
	A. Labor share			
lnauto95	-2.793**	0.0000251	-2.235**	-0.280
	(1.382)	(0.364)	(1.096)	(0.425)
N	454	408	408	408
	B. Wage			
lnauto95	0.0327	0.0140	0.0116	0.0169
	(0.0234)	(0.0325)	(0.0231)	(0.0220)
N	408	408	408	408
	C. (Share of) hours worked			
lnauto95	-0.0510***	-0.0905**	0.0248	0.0861***
	(0.0136)	(0.0380)	(0.0233)	(0.0238)
N	456	408	408	408

Note: The independent variables are lagged by one year, and all regressions control for other machinery innovation, GDP gap, country and year fixed effects. In Panel B, the dependent variable is the logarithm of the wage (deflated by the Producer Price Index). In Panel C, the dependent variable is logarithm of total hours and log of the ratio of hours by skill group to total hours, respectively. Outcome variables are constructed using data from World Input-Output Database. Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.3.4 Robustness checks

In this section, we present the results of several robustness checks.

Timing. In our benchmark specifications, we assume that automation innovation affects the labor share in the next period. In Table 2.3.4, we explore alternative timing assumptions. The table shows the results of our benchmark specification for different lags and leads. In Columns (1)–(2), the regressors are lead, in Column (3) the regressors are current, and in Columns (4)–(6) regressors are lagged. In all specifications except when lead by one period, the coefficient on automation is significantly negative. This suggests that our results do not depend on the chosen lag structure.

Table 2.3.4: Lags and Leads

	CLS					
	Leads			Lags		
	2	1	0	1	2	3
	(1)	(2)	(3)	(4)	(5)	(6)
lnauto95	-1.795**	-1.593	-2.949***	-2.739***	-1.683***	-1.342**
	(0.853)	(0.974)	(0.822)	(0.718)	(0.635)	(0.642)
lnother95	0.0904	0.116	1.127	1.303	0.808	0.181
	(0.787)	(0.996)	(0.852)	(0.832)	(0.704)	(0.755)
GDP gap	-14.43	-5.690	10.51	37.84***	41.71***	23.06
	(13.99)	(12.49)	(11.65)	(8.676)	(11.32)	(15.39)
Fixed effects	C+Y	C+Y	C+Y	C+Y	C+Y	C+Y
Observations	458	455	454	429	403	377
Ng	29	29	29	29	29	29

Note: Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

One explanation for this may be, as hypothesized previously, that the relationship between automation and the labor share are due to long-term processes. In Table 2.3.5 we explore the timing of the effects in our long-difference panel. Column (2) is the benchmark long-difference regression. In Column (1) and (3), we regress the 5-year change in the labor share on the change in the logarithm of automation innovation of the subsequent and preceding, respectively, 5-year interval. The estimated effects of growth in automation on the next and previous change to the labor share are insignificant. This suggests that for longer periods only changes in automation innovation only matters for the current labor share.

Table 2.3.5: Lags and Leads: Long Differences

		CLS	
	Lead	Current	Lag
	1	0	1
Δlnauto95	0.545	-4.026***	-0.406
	(2.329)	(1.508)	(1.100)
Δlnother95	-2.667	1.929	0.751
	(1.893)	(1.320)	(0.728)
ΔGDP gap	-32.48	74.99**	-48.45*
	(39.21)	(30.48)	(27.44)
Fixed effects	Y	Y	Y
Observations	52	75	50
Countries	28	28	27

Note: Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Alternative automation threshold. How robust are our results to variations in the definition of automation patents? We address this question using the laxer auto90 measure of automation, of which auto95 patents are a proper subset. In our sample, there are 73 percent more auto90 patents than auto95 patents. In Table 2.3.6, we repeat the baseline regressions using the laxer auto90 measure of automation innovation. The results are very similar to our benchmark regressions. As soon as non-automation machinery innovation is added as a control, the magnitude of the coefficient of these measures are equal.

Table 2.3.6: Baseline results with auto90

	Corporate Labor Share				
	(1)	(2)	(3)	(4)	(5)
lnauto90	-2.120***	-1.289***	-3.390***	-3.014**	-2.939**
	(0.399)	(0.474)	(1.260)	(1.238)	(1.254)
lnother90			2.188	1.798	0.680
			(1.349)	(1.288)	(2.528)
GDP gap				40.77***	40.55***
				(9.225)	(9.153)
lnnonmach					1.077
					(1.692)
Fixed effects	C	C+Y	C+Y	C+Y	C+Y
Observations	431	431	431	431	431
Countries	29	29	29	29	29

Note: Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Country-level clustering. In our benchmark regressions we use the Newey-West correction to estimate standard errors that are robust to serial correlation of order one and heteroskedasticity. In Table 2.5.2 we repeat our baseline tables with standard errors clustered at the country level. While standard errors are somewhat higher, all our results remain significant at the 5 percent significance level. Similarly, in Table 2.5.3, we show that our long-difference regressions results remain significant when clustering at the country level.

Alternative samples. We limit our sample to OECD countries, for which we think that automation patents is a good measure of automation. However, our results remain significant for larger samples. In Table 2.5.4, we run our benchmark specification on several samples. In Column (1)–(3) we restrict on OECD countries, whereas in column (4)–(5) we include all countries, for which we have data on the corporate labor share. In Column (1) and (4) we report the results of the benchmark specification. The coefficient of automation is significantly negative but of lower magnitude for the extended sample than for OECD countries. In Column (2) and (5) we restrict the sample on observations with at least ten auto95 patents. This ameliorates the issue of a low signal-to-noise ratio if the number of auto95 patents is very small. Under this restriction, the results become substantially stronger for the

OECD countries. Moreover, the coefficients for the extended sample are now highly significant and of similar magnitude as OECD countries. In Column (3) and (6) we weight countries by average real GDP in 1995–2011. The results become more pronounced for OECD countries and less pronounced on the extended sample.

Alternative patent quality restrictions. For our benchmark regressions we construct patent counts using biadic patents. This is to ensure that we only consider patents of sufficiently high quality. In Table 2.5.5, we report the regression results of our benchmark specification for several alternative ways to impose quality restrictions. In Column (1), we only count patents of innovations which have been patented in at least two of the following three patent offices: EPO, USPTO, or JPO. In Column (2), we count triadic innovations, which are inventions that have been patented at all of the three patent offices above. In Column (3), we count all patents which have received at least one forward citation. Finally, in Column (4), we consider all patents but weight them by the number of forward citations received until at most 5 years after publication. In all our specifications, except in Column (4), our estimates remain significantly negative and of similar magnitude. We are not concerned about the lack of association for the citations weighted patents. Citations are generally considered correlated with the market value of a patent. However, for instance, Abrams et al. (2013) have shown that this relationship may be U-shaped. Moreover, it may simply be the case that the market value of a patent is not strongly related the inventions' effect on the labor share.

Alternative sources for the labor share. In Table, 2.5.6 we use data from EU KLEMS and Penn World Table to construct measures of the labor share in the total economy rather than limited to the corporate sector. The results confirm the strong and significant negative results obtained for the corporate labor share.

2.4 Conclusion

In this paper, we document a robust negative relationship between automation and the labor share. The decline is mostly driven by the middle-skilled labor share.

We have shown that automation innovation is negatively associated with the share of hours worked of low-skilled labor, not significantly related to the share of hours worked of middle-skilled labor, and positively related to the share of hours worked of high-skilled labor.

Our results suggests that automation may play a significant role in the global decline of the share of value added paid out to labor. We are cautious about a causal interpretation of our results. Further research is necessary to establish the causality of the relationship we have documented in this paper.

2.5 Appendix

Table 2.5.1: Summary statistics

	N	Mean	Std. Dev.	p5	p50	p95
CLS	472	56.16	8.95	40.12	58.62	67.36
auto90	836	666.13	1223.70	0	145	3174
auto95	836	385.50	747.58	0	77	1937
other95	836	3103.78	4878.07	5	950	14254

Note: This table shows selected statistics on our patent counts and on the labor share. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5.2: Baseline regressions with country-level clustering

	Corporate Labor Share				
	(1)	(2)	(3)	(4)	(5)
lnauto95	-2.292***	-1.646**	-3.124**	-2.739**	-2.702**
	(0.573)	(0.679)	(1.181)	(1.114)	(1.068)
lnother95			1.649	1.303	1.037
			(0.991)	(0.865)	(2.135)
gdpgap				37.84***	37.89***
				(7.851)	(7.897)
lnnonmach					0.247
					(2.170)
Fixed effects	C	C+Y	C+Y	C+Y	C+Y
Observations	429	429	429	429	429
Countries	29	29	29	29	29

Note: Standard errors in brackets are clustered at the country-level. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5.3: Long Differences with country-level clustering

	5-year Changes ΔCLS			
	(1)	(2)	(3)	(4)
Δlnauto95	-1.754**	-1.431	-3.655**	-4.026**
	(0.680)	(1.023)	(1.679)	(1.581)
Δlnother95			2.341	1.929
			(1.615)	(1.460)
ΔGDP gap				74.99**
				(32.02)
Fixed effects	Y	Y	Y	Y
Observations	75	75	75	75
Countries	28	28	28	28

Note: Standard errors in brackets are clustered at the country-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5.4: Alternative samples

	Corporate Labor Share					
	OECD countries			All countries		
	(1)	(2)	(3)	(4)	(5)	(6)
lnauto95	-2.739***	-4.284***	-5.627***	-1.099**	-4.591***	-3.255*
	(0.718)	(1.036)	(1.509)	(0.455)	(0.979)	(1.704)
lnother95	1.303	0.699	0.232	0.0351	2.072*	-0.426
	(0.832)	(1.267)	(1.711)	(0.528)	(1.214)	(1.937)
gdpgap	37.84***	24.84**	17.11	22.29**	13.12	-5.745
	(8.676)	(10.38)	(16.07)	(9.122)	(9.582)	(21.03)
Fixed effects	C+Y	C+Y	C+Y	C+Y	C+Y	C+Y
$\geq$ 10 auto95 patents	No	Yes	Yes	No	Yes	Yes
GDP Weighted	No	No	Yes	No	No	Yes
Observations	429	394	394	604	481	481
Countries	29	29	29	47	41	41

Note: This table shows the baseline regressions with alternative sample restrictions. Columns (1) to (3) restricts the sample to the set of OECD countries. Column (1) is the baseline, column (2) restricts on observations with at least 10 auto95 patents, and column (3) weights observations by real GDP. Columns (4) to (6) show the same regression for all available countries. Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5.5: Alternative innovation quality restrictions

	Corporate Labor Share			
	Biadic (EP,JP,US) (1)	Triadic (2)	At least one citation (3)	Citations weighted (4)
lnauto95	-2.208***	-2.286***	-2.410***	-0.342
	(0.652)	(0.586)	(0.639)	(0.284)
lnother95	0.794	0.933	0.966	-0.158
	(0.721)	(0.642)	(0.705)	(0.504)
GDP gap	39.39***	42.59***	37.75***	42.03***
	(9.032)	(9.316)	(8.992)	(9.670)
Fixed effects	C	C+Y	C+Y	C+Y
Observations	429	429	429	429
Countries	29	29	29	29

Note: Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5.6: Alternative sources for the labor share

	Labor Share			
	(1)	(2)	(3)	(4)
		A. EU KLEMS		
lnauto95	-0.423*	-0.607**	-1.766***	-1.329***
	(0.257)	(0.281)	(0.487)	(0.483)
lnother95			1.398**	0.869
			(0.570)	(0.544)
GDP gap				46.53***
				(9.779)
Observations	434	434	434	434
Countries	24	24	24	24
		B. Penn World Table		
lnauto95	-1.465***	-0.519	-1.931***	-1.654***
	(0.291)	(0.318)	(0.505)	(0.490)
lnother95			1.592***	1.274**
			(0.572)	(0.546)
GDP gap				33.21***
				(8.892)
Observations	627	627	627	626
Countries	34	34	34	34
Fixed effects	C	C+Y	C+Y	C+Y

Note: In both panels, the outcome variable is the total economy labor share. Standard errors in brackets are robust to heteroskedasticity and first-order serial correlation using Newey-West. The independent variables are lagged by one year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

Opportunity and Inequality Across Generations

This chapter is joint work with Winfried Koeniger. *

A version of this paper is under review at the Journal of the European Economic Association.

3.1 Introduction

Inequality across generations is transmitted through parents' nurture and nature. A central question is how much the opportunities of each generation depend on nurture and nature. A tightly related question is whether, and how, optimal policy should change the observed patterns.

To shed light on these issues, we analyze inequality and opportunity across generations in an economy with family dynasties calibrated to the U.S. The transmission of inequality across generations in the model is influenced by nurture, in terms of bequests and schooling investment, and nature, in terms of ability that is private information, stochastic and persistent across generations. Mobility within a

*Winfried Koeniger: Department of Economics, University of St.Gallen, SEW-HSG, Varnbüelstrasse 14, 9000 St.Gallen, Switzerland, winfried.koeniger@unisg.ch. Carlo Zanella: Department of Economics, University of Zurich, Schönberggasse 1, 8001 Zurich, Switzerland, carlo.zanella@econ.uzh.ch. We thank Julien Prat for discussions initiating this project and participants at various seminars for very helpful comments.

generation, resulting from the stochastic changes in ability, indicates the extent of opportunities available to each generation.

Based on the model, we make three contributions. Firstly, we use the calibrated model to complement the vast empirical literature on inequality and mobility by illustrating mechanisms through which nurture and nature affect mobility and the intergenerational transmission of inequality. Secondly, we analyze how mobility and the transmission of inequality across generations would change if, starting from the calibrated steady state, a reform implemented the social optimum. Thirdly, we compare welfare of economies with simple tax and subsidy systems to welfare in the social optimum, which, in theory, requires complex history-dependent tax schedules for its implementation as in Farhi and Werning (2010). To the best of our knowledge, such analysis of simple tax systems has not been provided yet in the context of intergenerational models with persistence in unobservable ability. We now elaborate further on each of the three contributions.

Concerning the intergenerational transmission of inequality, we find important differences in how bequests and schooling affect this transmission in our calibrated model. Bequests or inheritances decrease the incentive to exert labor effort through a negative wealth effect and thus induce mean reversion in labor earnings.[1] Investment into schooling instead increases the persistence in labor earnings across generations because more human capital increases labor productivity and thus also labor effort, as long as the substitution effect of the productivity increase dominates. Quantitatively, we find that the transmission of inequality is mostly determined by nature (ability) rather than nurture (bequests or human capital investment), in line with recent empirical findings of Bingley et al. (2018) who identify the effects of nature and nurture with a credible "Children of Twins" design.[2]

[1]Such a negative wealth effect on life-cycle labor supply, through early retirement or less labor force participation of some household members, is supported empirically by Holtz-Eakin et al. (1993) and Brown et al. (2010). See also the evidence for lottery winners by Imbens et al. (2001) for the U.S., by Cesarini et al. (2017) for Sweden, and the analysis of Kindermann et al. (2018) on the consequences of this wealth effect for labor income taxation. The wealth effect also implies less investment into schooling because schooling results in more resources only if combined with labor effort. Although the negative wealth effect dominates on average, we find that more wealth relaxes borrowing constraints for some dynasties and thus increases their investment into schooling.

[2]The empirical importance of nature and nurture for the intergenerational transmission of inequality is still a matter of debate. Lee and Seshadri (2019) argue, based on a rich structural model, that the empirical estimates may overstate the importance of nature because they do not

Our analysis of the intergenerational transmission of inequality and mobility in the model calibrated to the U.S. complements the empirical evidence provided by Chetty et al. (2014) for the U.S.[3] We highlight economic mechanisms, as discussed above, that generate the empirically observed patterns and we quantify the trade-off between intergenerational insurance and mobility. In the analysis of the reform that implements the social optimum, more income inequality does not make dynasties worse off *ex ante* because the reform also increases intergenerational insurance.

The importance of initial conditions for the welfare of generations has been emphasized by Keane and Wolpin (1997) and more recently by Huggett et al. (2011), De Nardi and Yang (2016) and Lee and Seshadri (2019). Structurally estimating career decisions in the U.S., Keane and Wolpin (1997) find that initial conditions at age 16 explain 90 percent of the total variance in expected lifetime utility. Based on a model with risky human capital, Huggett et al. (2011) find that differences in initial conditions at labor market entry account for more than 60% of the variation in lifetime utility. This percentage increases to more than 70% in Lee and Seshadri (2019) who model human capital formation early in life. The dominance of initial conditions for each generation's welfare motivates our focus on the opportunities and the transmission of inequality across generations rather than on differences that arise due to shocks within the labor-market career of a generation.

In their intergenerational model with estate taxation calibrated to the U.S., De Nardi and Yang (2016) find that parental background matters most for life-time utility at the top of the distribution of parental earnings, given that the calibrated distribution is very unequal at the top. The difference of being born into a family in the lowest earnings state compared to the second-lowest earnings state is very small instead. We complement this evidence by showing that family background in terms of parent's schooling investment is most effective at the top of the ability distribution because of the complementarity of ability and schooling. At the bottom of the ability distribution, bequests are much more effective in welfare terms instead.

We characterize the social optimum as the solution of a dynamic Mirrleesian problem, in which asymmetric information constrains the insurance that the planner

account for general equilibrium effects.

[3] Güell et al. (2018) provide evidence for Italy. See also their references for further literature.

can provide against the ability risk that dynasties face. This makes the quantitative comparison between the calibrated economy and the social optimum non-trivial. Inequality and less than full insurance are also a feature of the social optimum, which provides a useful benchmark to put the observed mobility and transmission of inequality in perspective.

Our analysis of the social optimum builds on Phelan (2006) and Farhi and Werning (2007) who have shown that the social optimum in a dynamic Mirrleesian economy with asymmetric information need not imply immiseration as in Atkeson and Lucas (1992) if the planner discounts the future but attaches more weight to future generations than implied by the altruism of a family dynasty. We proceed as in Farhi and Werning (2007) and assume that dynasties are weighed equally in the planner's problem. This implies a wedge between the discount rate that the planner and the family dynasty apply to the utility of each generation: the planner cares directly about the welfare of a future generation and also indirectly, given that family dynasties care about their offspring. The non-degenerate steady-state distribution in the social optimum, resulting from the wedge in discount rates, allows for a meaningful analysis of the transition from the steady state of the calibrated economy to the social optimum. Because the wedge between the discount rates is an important determinant of the distribution in the social optimum, we explain in Section 3.5 how it relates to the difference between the real interest rate and the discount rate of the family dynasty if we focus on the case with a stable consumption distribution in the long run. We discipline this difference in the calibration by matching median bequests observed in U.S. data.

In our quantitative analysis of the reform and the subsequent transition to the social optimum, we find that the importance of nurture (in terms of bequests and human capital investment) increases relative to nature (in terms of ability) compared with the steady state of the calibrated economy. We show that this implies more intergenerational insurance against ability risk at the cost of less social mobility, in the sense that ability is less important for the rank in the welfare distribution within each generation. We quantify this trade-off between insurance across generations and social mobility in the transition to the social optimum, in which the extent of intergenerational insurance and mobility is determined by optimal incentive provision as pointed out by Phelan (2006).

We find that production is decoupled more from consumption after the reform on the transition to the social optimum. The pass-through coefficient of an unexpected ability shock to consumption decreases by more than half implying more intergenerational insurance. At the same time, the correlation of ability and labor effort increases in the social optimum compared to the steady state calibrated to the U.S., and income mobility across and within generations remains quite stable after the reform on the transition to the social optimum.

At the time of the reform, we hold constant the present discounted value of the expected net costs for the allocation of each dynasty and give equal weight to dynasties in the planner's objective function. We thus focus on insurance and refrain from adding motives for redistribution across dynasties by applying different weights in the objective function. Our reform experiment thus answers the question how much additional insurance the planner provides ex post by using an optimal policy without redistributing across dynasties at the time of the reform. This makes the reform also implementable from a politico-economic point of view.

Our analysis illustrates the general point, made in Lee and Seshadri (2018) for example, that trade-offs arise depending on whether equality of opportunity is considered from a dynastic point of view, tilting the balance towards more insurance across generations at the cost of less social mobility within generations, or from the point of view of individual families within a generation, tilting the balance towards more mobility within a generation and thus less intergenerational insurance.[4] Interpretation of the empirical evidence on mobility patterns and inequality provided in Chetty et al. (2014) for the U.S., Güell et al. (2018) for Italy or Adermon et al. (2018) for Sweden, for example, thus requires assumptions about the social welfare function. The welfare function used in this paper contains a dynastic motive which is disciplined by calibrating the discount factor of dynasties to match the empirically observed size of bequests.

Concerning the comparison of economies with simple tax and subsidy systems to the social optimum, we find that about half of the welfare gains of moving from the laissez faire to the social optimum can be achieved in economies with linear taxes and

[4]See Arneson (2018) for a recent discussion of different interpretations of equal opportunity.

subsidies that condition on current labor income, bequests and schooling. We find that bequests and schooling are both subsidized but at different rates. This relates to results in Farhi and Werning (2010) who show that, in an implementation with history-dependent tax schedules, bequests and human capital should be subsidized. In our model, subsidies for schooling investment and bequests play a different role in shaping inequality and opportunity also because we relax the assumptions in Farhi and Werning (2010) that children make no labor supply decision and that there is no uncertainty.[5] In the tax and subsidy system with optimal linear rates, 18% of labor income is taxed and bequests and schooling are subsidized at rates of 36% and 30%, respectively.

Our results show to which extent the endogenous state variables bequests and human capital together with labor income, which depends on ability draw of the current generation, can capture the effect of history in our calibrated economy with persistent ability shocks. Albanesi and Sleet (2006) have shown that history can be fully summarized by conditioning optimal taxes on the endogenous state variable if shocks to unobserved ability are i.i.d. and thus not persistent.[6] In our model with persistent shocks to ability, history-independent tax schedules that condition on bequests, schooling and labor income only allow to implement the social optimum approximately.[7]

Our analysis also relates to the large strand of literature on optimal taxation of human capital or bequests. The special issue on human capital and inequality edited by Corbae et al. (2017) and the volume on inequality and redistribution of the Carnegie-Rochester-Conference (2016) provide a good overview over recent research. Optimal taxation of human capital using a Mirrleesian approach has been analyzed by Findeisen and Sachs (2016), Kapička (2015), Kapička and Neira (2019), Stantcheva (2015) and Stantcheva (2017), and Koeniger and Prat (2018). Heathcote et al. (2017), Krueger and Ludwig (2016), Lee and Seshadri (2019) and Peterman (2016) are examples for analyses based on a Ramsey approach to optimal

[5]Farhi and Werning (2010) mention in their discussion of proposition 5 for history-dependent tax schedules that the symmetry in the taxation of bequests and human capital only holds under these assumptions.

[6]Stantcheva (2017) has extended this result in a life-cycle model with assets and human capital.

[7]Given that shocks to ability are unobservable and persistent, our recursive formulation of the planner problem relies on results of Kapička (2013) and Pavan et al. (2014).

taxation. Farhi and Werning (2010) and Phelan and Rustichini (2018) analyze optimal taxation of bequests or inheritances. The approach in our paper is to compare the status quo in the U.S., using approximations of existing tax schedules, with a social optimum that does not imply immiseration in the long run, and to analyze approximations of the social optimum with simple history-*in*dependent tax schedules. The focus of this paper is thus different from Koeniger and Prat (2018), who characterize socially optimal taxation of bequests and human capital if the planner and the dynasties apply the same weight to future generations, implying immiseration in the long run as in Atkeson and Lucas (1992).

Our analysis proceeds in the following steps. In Section 3.2, we model the decision problem of family dynasties. We then calibrate the model to U.S. data in Section 3.3. In Section 3.4, we provide results for inequality, mobility and the importance of parental background in this economy. To benchmark these results, we characterize the planner problem in Section 3.5. We compute the constrained-efficient social optimum and analyze in Section 3.6 how a tax reform changes inequality and mobility on the transition to the socially optimal steady state. In Section 3.7, we compare economies with simple history-independent tax schedules to the social optimum. We discuss the properties of these schedules and conclude in Section 3.8. The appendix contains a robustness analysis for alternative assumptions on the complementarity of schooling and ability in the production function, on the persistence of ability, on the Frisch elasticity of labor supply, and on the bequest target in the calibration.

3.2 The model

We build on the dynasty model of Koeniger and Prat (2018) to understand key mechanisms through which nurture, in terms of schooling and bequests, and nature, in terms of ability, affect inequality and mobility across generations. We analyze decisions of family dynasties who are composed of parents and children in each generation, have an infinite planning horizon and a size normalized to one. Each generation of a dynasty chooses, conditional on the parents' ability draw, the labor supply of the parents, consumption, and the bequests and schooling for the children.

Our dynasty model is deliberately simpler than the model by Lee and Seshadri (2019) who analyze the sequence of decisions over the life cycle for each generation in more detail. The simplicity of our model keeps the problem tractable when we solve for the social optimum. This allows us make two of our contributions in this paper: to analyze how the transmission of intergenerational inequality changes on the transition from the calibrated steady state to the social optimum; and to characterize simple tax schedules which implement the social optimum approximately.

Preferences are time separable across generations and we make the common assumption that the per-period utility function $\mathbf{U}(c_t, l_t)$ is separable in consumption c_t and labor effort l_t:

$$\mathbf{U}(c_t, l_t) = u(c_t) - \mathbf{v}(l_t),$$

where $u(c_t) \in \mathcal{C}^2(\mathbb{R}_+)$ is increasing in c_t and strictly concave, and $\mathbf{v}(l_t) \in \mathcal{C}^2(\mathbb{R}_+)$ is increasing in l_t and strictly convex.

Each generation of a family differs in its ability θ_t. Ability is not observable so that tax schedules cannot be conditioned on it. Bequests b_t and human capital h_t, think of years of schooling and high-school or college degrees, are public knowledge instead. Output y_t is produced with the technology $Y(h_t, l_t, \theta_t)$ which is increasing in its arguments and concave. Although output y_t is observable, actual labor supply l_t cannot be inferred from it because ability θ_t is stochastic and hidden.

The expenditure of schooling $g(h_{t+1}, h_t)$ depends on the amount of human capital investment h_{t+1} into the children and on the family background, which we summarize by the stock of human capital of parents h_t. This cost function follows from inverting a human-capital production function in the spirit of Ben-Porath (1967), where human capital of the next generation depends on the expenditure on schooling and parental background.[8]

At the beginning of each period, the dynasty learns the ability of the parents and then makes its choices about labor supply, consumption, bequests and human

[8] We abstract from modeling a parental time input because such an input is plausibly unobservable which would render the analysis much less tractable. We also abstract from a direct influence of the childrens' ability on the cost of human capital investment for parsimony. This would add another channel through which output would depend on ability but would not add further insights as long as the observation of human capital investments does not provide information about ability.

capital investment. Ability is drawn from the bounded interval $\Theta \equiv [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$, where we assume a continuously differentiable distribution $F : \Theta \to [0,1]$ with conditional density $f(\theta_t | \theta_{t-1})$. For the analysis of the planner's problem we make the further assumptions that $f(\theta_t | \theta_{t-1})$ has full support, that it is of class $\mathcal{C}^2$ with respect to its second argument θ_{t-1}, and that it has a bounded derivative $|\partial f(\theta_t | \theta_{t-1}) / \partial \theta_{t-1}| \leq B$ for some $B \in \mathbb{R}_+$. The dependence of the distribution $F(\theta_t | \theta_{t-1})$ on the type of the previous generation allows us to model intergenerational transmission of ability, which may occur because of genetic inheritance or nurture in early childhood. When we quantify the effect of nature in this paper, the effect should thus be considered an upper bound, given that ability may also contain some nurture component. In the calibration discussed in Section 3.3, we discipline the transmission of ability by choosing the correlation between θ_t and θ_{t-1} such that the intergenerational correlation of earnings based on model simulations matches the empirical counterpart in U.S. data.

The stationary recursive problem of the family dynasty is

$$\begin{aligned} \widehat{W}(b,h,\theta) = & \max_{\{b',h',l,c\}} \left\{ \mathbf{U}(c,l) + \beta \int_{\Theta} \widehat{W}(b',h',\theta')\, dF(\theta'|\theta) \right\} \qquad (3.1) \\ \text{s.t. } b' = & (1+r)b - T^b(b) - c - g(h',h) - T^h(h',h) + y - T^y(y), \\ b' \geq & \max\{-\phi g(h',h), \underline{b'}\}, \\ y = & Y(h,\theta,l), \\ \ln(\theta') = & \rho \ln(\theta) + \epsilon, \end{aligned}$$

where β is the discount factor of the family measuring the strength of the altruism towards future generations, $0 \leq \rho \leq 1$ captures the persistence of shocks to ability, and "′" denotes values of variables one period in the future. Families can pass on the fraction $\phi \in (0;1)$ of the schooling expenditures to the next generation but the borrowing constraint limits the overall student debt to $\underline{b'}$. The functions $T^i(\cdot)$, for $i = b,h,y$, denote the non-linear schedules for taxes and subsidies on bequests, education and labor income, respectively. Since $g(h',h)$ and $T^h(h',h)$ enter additively in the budget constraint, we set $T^h(h',h) = 0$ and interpret $g(h',h)$ as the *net* cost of human capital accumulation.

3.3 Calibration

In this section, we calibrate the model of family dynasties to the U.S. economy. We will use the calibrated economy to characterize inequality, mobility and the effect of parental background in Section 3.4. In Section 3.6 we then use the steady state of the calibrate economy as a starting point to analyze the effect of a tax reform on inequality and mobility, on the transition to the socially-optimal steady state.

Preferences.—A key parameter in our calibration is the wedge between the rate at which the planner and the dynasty discount the welfare of future generations. Although this may seem a minor detail, it has major implications for the stationary distribution in the social optimum. As we are going to see in Section 3.5, a stationary distribution in the planner problem requires that the discount rate of the planner equals the intertemporal marginal rate of transformation, i.e., the real interest rate. It is well known that in the calibrated economy, in which dynasties face incomplete markets, stationarity requires that the discount rate of the families is higher than the real interest rate so that the discount factor $\beta < 1/(1+r)$. Thus, we discipline the wedge in the discount rate between the families and the planner by calibrating β so that, for a given real interest rate of 3%, the stationary distribution for bequests in the model implies a median for bequests that matches the median observed in the data, conditional on receiving a positive bequest.[9]

Evidence in table 2 of Wolff and Gittleman (2014), based on the Survey of Consumer Finances (SCF) in the time period $1989-2007$, shows that the median wealth transfer among households in the U.S. has been $\$73,600$, conditional on receiving a transfer.[10] We adjust this figure for household size dividing it by 1.4.[11]

The rest of our calibration strategy follows Koeniger and Prat (2018) by and

[9]In the robustness analysis discussed in Section 3.6 we report results if we match the conditional *mean* of bequests instead.

[10]The value is expressed in terms of dollars in 2007. Table 1 in Wolff and Gittleman (2014) shows that 84% of the wealth transfers are classified as inheritances. Most other transfers are classified as gifts and most transfers are from family members. Given that the timing of intergenerational transfers before or after death is difficult to map into our model, we consider all wealth transfers.

[11]This number is reported in table 8 of Hintermaier and Koeniger (2011) who compute household size for the waves of the SCF in the same time period, based on an equivalence scale that assigns a weight of 1 to the first person in the household, 0.34 to the second person and approximately 0.3 to each additional member of the household. See also Fernández-Villaverde and Krueger (2007), table 1, last column.

large and we repeat it here for completeness. We specify the utility function as $\mathbf{U}(c,l) = \ln(c) - l^{\alpha}/\alpha$, which satisfies the assumptions for the utility function made in Section 3.2. The estimate for the Frisch elasticity of 0.5 documented in Chetty (2012) implies that $\alpha = \varepsilon^{-1} + 1 = 3$.

Technology.—The length of a period in the model is 30 years to approximate the time until labor-market entry of a new-born generation and the length of the labor-market career. We set the annualized real interest rate to 3% and assume the production technology

$$Y\left(h, l, \theta\right) = A\left(\theta, h\right) l$$

with labor productivity

$$A\left(\theta, h\right) = \left[\xi \theta^{\frac{\chi-1}{\chi}} + (1-\xi) h^{\frac{\chi-1}{\chi}}\right]^{\frac{\chi}{\chi-1}}$$

and $\chi \in [0, \infty)$, $\xi \in (0, 1)$.

The linearity of the production technology in labor effort and the assumption of a given interest rate, which is not influenced by accumulation behavior within the U.S., imply that we can solve the problem of the dynasties separately from each other. As a benchmark, we assume that the elasticity of substitution $\chi = 1$ so that labor productivity is a Cobb-Douglas function of ability and human capital: $A(\theta, h) = \theta^{\xi} h^{1-\xi}$. We will check the robustness of our results for a different degree of complementarity between ability and schooling.

Cobb-Douglas productivity has the advantage that, for competitive labor markets, wages $w(\theta, h)$ are log-linear in human capital and unobserved ability:

$$\ln w(\theta, h) = \ln A(\theta, h) = (1-\xi) \ln h + \xi \ln \theta, \tag{3.2}$$

so that it is straightforward to use the variance of residual wages as target to calibrate the variance of unobserved ability θ.[12] We assume that θ is drawn from a log-normal distribution. The mean and standard deviation specified in Table

[12] Given (3.2), the variance of residual wages is the variance of wages which remains after regressing log-wages on years of schooling where, in our model, chosen years of schooling S beyond the compulsory school years correspond to $\ln h$.

3.3.1 imply a variance of residual wages of 0.2. This corresponds to estimates by Heathcote et al. (2010) for the U.S. in 2005, for the part of the variance of residual log-wages that is generated by persistent shocks.[13] We use the variance resulting from persistent shocks as target because a generation's labor-market career takes place within a period in our model so that transitory shocks (at least partially) cancel out and θ, within a period, is fully persistent.

We refer to the large empirical evidence on Mincerian wage regressions to calibrate the parameter ξ of the production function. In his survey, Card (1999) shows that the marginal return to schooling is quite robustly estimated across studies and close to 10%. Equation (3.2) thus implies $1-\xi = 0.1$, given that years of schooling S correspond to $\ln h$ in our model, where compulsory schooling is defined as $h = 1$ in which case the chosen years of schooling $S = \ln 1 = 0$.

Borrowing opportunities.—We set the parameters of the borrowing constraint in problem (3.1) to $\phi = 0.5$ and $\underline{b'} = -\$30,000$. This implies that families can finance up to 50% of their human capital investment into their children, with a maximal amount of debt of $\$30,000$. At the time the next generation makes its choices the accrued interest then implies a maximal total debt of $\$72,818$ so that the amount for outstanding student loans broadly matches the amounts reported in Lee et al. (2014).

Approximation of tax schedules.—We use the parametrization proposed by Heathcote et al. (2017) to approximate labor income taxes in the U.S.: $T^y(y) = y - \delta y^{1-t_y}$, with $t_y = 0.181$ and $\delta = 0.9276$.[14] We approximate taxes on bequests using the parametrization for estate taxes proposed by De Nardi and Yang (2016) because our model does not distinguish estates from bequests. Thus, families pay 20% tax if the bequest exceeds the exemption of $\$756,000$. The function $g(h', h)$ captures net education costs after subsidies and we now discuss the calibration of its parameters.

Stochastic process for ability and education costs.—The parameters for the per-

[13] See panel C of Figure 3 in Heathcote et al. (2008).

[14] $T^y(y)$ is negative if $y < \delta^{\frac{1}{t_y}} \approx 2/3$. A unit in our model corresponds to mean earnings of high-school dropouts, as explained further below. Thus, in our calibrated model, workers receive positive transfers if their annual income is below $\$14,423$. Otherwise they pay taxes.

Parameters	Target	Source
Preferences		
Discount factor (annualized): $\beta = 0.966$	Median bequest (conditional on receiving one, equivalized): \$52,571	Wolff and Gittleman (2014)
Disutility of labour $v(l) = l^{\alpha}/\alpha$: $\alpha = 3$	Frisch elasticity: 1/2	Chetty (2012)
Storage technology		
$r = 0.03$	Annualized real interest rate	Standard
Production technology: $y/l = \theta^{\xi} h^{1-\xi}$		
$\xi = 0.9$	Returns to education: 10%	Card (1999)
Borrowing opportunities		
$\phi = 0.5$, $\underline{b'} = -\$30,000$	Student loans in FRBNY Consumer Credit Panel	Lee et al. (2014)
Taxes		
$T^y(y) = y - \delta y^{1-t_y}$, $t_y = 0.181$, $\delta = 0.9276$	Parametric approximation of the U.S. labor income tax schedule	Heathcote et al. (2017)
$T^b(b) = \max\{t_b(b - x_b), 0\}$, $t_b = 0.2$, $x_b = \$756{,}000 \times$ model-unit factor	Parametric approximation of the U.S. estate tax schedule	De Nardi and Yang (2016)
AR(1)-process for ability: $\ln(\theta') = \rho \ln(\theta) + \epsilon$, $\ln\theta' \sim \mathcal{N}\left(-\frac{\sigma_\epsilon^2}{2(1-\rho^2)}, \frac{\sigma_\epsilon^2}{1-\rho^2}\right)$		
$\rho = 0.448$	Intergenerational earnings elasticity: $\iota = 0.45$	Chetty et al. (2014)
$\frac{\sigma_\epsilon^2}{1-\rho^2} = \frac{0.2}{\xi^2}$	Variance of residual wages: 0.2	Heathcote et al. (2008)
Education cost: $g(h', h) = \kappa (h')^{\varsigma_1} h^{\varsigma_2}$		
$\kappa = 0.0014$	Average years of schooling: 12.86	Barro and Lee (2013)
$\varsigma_1 = 0.7465$	Average net cost for an additional year of education: \$13,845	OECD (2011), Stantcheva (2017)
$\varsigma_2 = -0.0005$	Intergenerational correlation of years of schooling: 0.46	Hertz et al. (2008)

Table 3.3.1: Calibrated parameter values

sistence of ability shocks and the education cost function are calibrated jointly together with the discount factor to match the following target statistics: median bequests, the average years of schooling, the average net cost of an additional year of secondary/tertiary education, the correlation between years of schooling across generations and the intergenerational earnings elasticity.

We have chosen the target moments so that they are tightly related to the parameters we calibrate. Although jointly calibrated, each target is closely related to the calibration of one of the parameters. As mentioned above, calibration of the discount factor β helps to match the median bequest. The persistence in the stochastic process of ability allows us to match the intergenerational earnings elasticity ι (IGE), resulting from a linear regression of $\ln y'$ on $\ln y$. This is intuitive because ability θ affects labor earnings through changes in labor productivity $A(\theta, h)$. Given that the exponent of ability ξ in the Cobb-Douglas function for labor productivity is nine times higher than the exponent of human capital, the IGE is mostly determined by the persistence of ability that is fed into the model. The endogenous choices of human capital and labor supply quantitatively matter much less for labor earnings and thus for the IGE.

The parameters κ and ς_1 of the cost function $g(h', h)$ in Table 3.3.1 allow us to match average years of schooling and average net cost for schooling. The annual expenditure per student and year in the U.S. is \$12,690 for upper-secondary education and \$29,910 for tertiary education, as documented in tables B.1.2 and B.1.6 of OECD (2011). The average cost for an additional year of schooling is thus \$21,300. We assume, as in Stantcheva (2017), that 35% of expenses for human capital investment related to higher education are subsidized so that we get a target of \$13,845 for the cost net of the subsidy for a student at the time of high-school graduation.

We have to convert the monetary costs observed in the data into units of the model. We make the empirically plausible assumption that the average family *without* any non-compulsory education does not receive, or leave, any bequests and does not spend significant amounts on education. Such a family generation then approximately consumes all resources in a hand-to-mouth fashion so that income per model period corresponds to 0.9356 in model units.[15] Expressed in dollars, this amount equals the mean annual earnings of high-school dropouts of \$20,241 in 2010 which have a present value for a 30-year period of \$436,762.[16]

The calibrated value of ς_1 implies sufficient convexity of the cost function so that the calibrated economy or the planner problem analyzed in Sections 3.5 and 3.6 are concave.[17] The calibrated parameter ς_2 of the cost function is close to zero. This implies that the model matches the intergenerational correlation in the years of schooling although parental background reduces the net cost of education only very mildly: if parents have four years of non-compulsory schooling, this reduces the cost of educating their children only by 2 per mille.

Simulation.—We solve the problem by applying the endogenous gridpoint method

[15] A hand-to-mouth consumer without bequests consumes net income, $c = y - T^y(y) = \delta y^{1-t_y}$ where the last step follows using the tax schedule $T^y(y) = y - \delta y^{1-t_y}$. The optimal labor supply for a hand-to-mouth consumer without bequests is $l^*(\theta, h) \equiv \arg\max \left\{ \ln(\delta \left[A(\theta, h)\, l\right]^{1-t_y}) - v(l) \right\}$. For $v(l) = l^\alpha/\alpha$, we obtain $l^*(\theta, h) = (1 - t_y)^{1/\alpha}$. $A(1,1) = 1$ then implies that the period income of the average worker without any non-compulsory education is $y^*(1,1) = (1 - t_y)^{1/\alpha} = 0.9356$, once we insert the parameter values documented in Table 3.3.1.

[16] See Table 232 in the *Statistical Abstract of the United States 2012* available at https://www.census.gov .

[17] In particular, $\varsigma_1 > 1 - \xi$ where $1 - \xi$ is the exponent of human capital in the production function. Concavity makes our problem tractable because otherwise we could no longer rely on the first-order approach to characterize the social optimum.

Variable	*Data* (1)	*Model* (2)
Median bequests, conditional on $b > 0$	\$52,571	\$52,717
Average years of schooling S	12.86	12.75
Correlation(S', S)	0.46	0.48
Intergenerational earnings elasticity	0.45	0.45
Average net cost of an additional year of schooling	$\$13,845$	$\$13,674$

Table 3.3.2: Target statistics in the data and model predictions

proposed in Hintermaier and Koeniger (2010). The algorithm is described further in the online Appendix C of Koeniger and Prat (2018). For the simulations, we draw $500,000$ observations, simulate the respective paths based on the model solution for 100 generations to obtain a stationary distribution. Further implementation details on the numerical solution and calibration are provided in Appendix 3.9.2.

Table 3.3.2 shows that the calibrated model matches the data targets quite closely. In Appendix 3.9.5 we provide a set of robustness checks. In these alternative calibrations we target the conditional mean instead of the conditional median of bequests, we target a lower intergenerational earnings elasticity, we calibrate a higher Frisch elasticity, and we allow for a higher complementarity between human capital and ability in the function for productivity.

3.4 Opportunity and inequality in the calibrated economy

We investigate the transmission of inequality across generations in the calibrated economy. In doing so, we highlight the mechanisms through which parental background affects this transmission, both in terms of nurture through bequests and human capital, and in terms of nature through innate ability.

Rank-rank correlations		
Across generations		
	Ability	0.43
	Welfare	0.67
	Labor earnings	0.43
	Consumption	0.64
	Bequests	0.59
	Years of schooling	0.46
	Labor effort	0.34
Within generation		
	Ability and Welfare	0.90

Elasticity of labor earnings with respect to...		
Nurture		
	Bequests	-0.033
	Schooling	0.50
Nature		
	Ability	0.93

Table 3.4.1: Mobility and the effect of nurture and nature on earnings

Notes: As discussed in Section 3.3, the empirical literature on the returns to education implies that the elasticities of wages (or productivity) with respect to ability and human capital in our model are 0.9 and 0.1, respectively. This gives empirical content to the changes in schooling and ability for the interpretation of the labor earnings elasticities.

The top panel of Table 3.4.1 shows rank-rank correlations across generations for all variables of interest in the steady state of the calibrated economy. For earnings, we can compare this with empirical results for the U.S. in Chetty et al. (2014) who report a rank-rank correlation of 0.34 in their table 1.[18] In our model it takes on average slightly more than five generations until the offspring of a family in the bottom decile of the income distribution reaches the mean income. This is in line with results in OECD (2018) for the U.S. that it takes on average four to five

[18]The estimated transition matrix reported in Table 2 of Chetty et al. (2014) is remarkably similar to Table 3.9.1 in Appendix 3.9.4, generated by our calibrated model. The matrix in Chetty et al. (2014) predicts somewhat less persistence at the bottom and top of the income distribution. As shown in Table 3.9.6 of Appendix 3.9.5, we match this matrix more closely in the robustness check for a calibration with less persistence that targets the lower end of estimates for the intergenerational earnings elasticity reported in Table 1 of Chetty et al. (2014). The model-implied rank-rank correlation of earnings is 0.29 in this case.

generations for the offspring of a low-income family to reach the average income.

The rank-rank correlations in the top panel of Table 3.4.1 reveal that all variables, but for labor effort, are at least as persistent across generations as ability, where the persistence of ability is given by the calibrated stochastic process. The difference between the rank-rank correlation of labor earnings and the respective correlations for consumption or welfare in Table 3.4.1 shows that labor earnings are more mobile (and hence less persistent) across generations than consumption or welfare. This is because dynasties partially self insure ability shocks and thus smooth consumption across generations, implying a less than perfect rank-rank correlation of 0.90 between ability and welfare within generations. The insurance against ability shocks is achieved by nurture which implies persistence in bequests and schooling. The relatively lower persistence of labor effort than ability is the result of different effects: more bequests result in a negative wealth effect on labor effort but more schooling or ability make labor effort more productive so that the substitution effect would increase labor effort, ceteris paribus. These effects also explain the different sign of the average steady-state elasticities of labor earnings with respect to changes in bequests, schooling and ability, respectively, reported in the bottom panel of Table 3.4.1.

	Quintiles				
Ability / Welfare	1	2	3	4	5
1	0.86	0.08	0.03	0.02	0.01
	(-6.03,4.4)	(37.64,4.8)	(93.7,4.8)	(161.32,4.8)	(283.21,4.8)
2	0.14	0.66	0.11	0.06	0.03
	(-8.1,4.4)	(-5.0,4.6)	(45.71,4.9)	(116.35,4.9)	(245.21,4.9)
3	0.0	0.26	0.53	0.14	0.07
	(.)	(-8.4,4.5)	(0.01,4.7)	(72.03,5.0)	(205.73,5.0)
4	0.0	0.0	0.33	0.5	0.17
	(.)	(.)	(-7.88,4.6)	(13.55,4.8)	(157.32,5.0)
5	0.0	0.0	0.0	0.28	0.72
	(.)	(.)	(.)	(-4.6,4.7)	(76.71,5.0)

Table 3.4.2: Social mobility matrix in the calibrated economy

Notes: Each cell contains the probability of a family in an ability quintile to be in a specific quintile of the families' welfare distribution. In brackets for each cell, we report the average values of the state variables other than ability. Bequests are in units of $1,000 and school years are non-compulsory. The probabilities across columns in each row may not add up to 1 because of rounding.

To investigate the extent of social mobility in the calibrated economy further, we show in Table 3.4.2 how shocks to ability translate into differences in welfare. The matrix in Table 3.4.2 displays the probability that a family in a quintile of the ability distribution is in a specific quintile of the welfare distribution. The welfare measure is dynastic and includes the discounted welfare of future generations. We define perfect social mobility within a generation as the case in which nature, in terms of ability θ, fully determines the position in the welfare distribution. In other words, nurture, in terms of received bequests b and obtained human-capital investment h, is then irrelevant for the position in the welfare distribution. In this case, the matrix would be an identity matrix. The more weight is on the off-diagonal elements, the more nurture dampens the effect of nature on the position in the welfare distribution and thus insures generations from an intergenerational perspective.[19]

For example, the cell in the first row and fifth column of the matrix shows that a family in the first quintile of the ability distribution has a one percent probability of being in the top quintile of the welfare distribution. These families have received $\$283,210$ as bequests and have obtained 4.8 years of non-compulsory schooling on average. This shows that nurture can compensate for bad draws of nature. Bequests are more effective in compensating for low ability than human capital investments because ability and schooling are complements in making labor effort more productive. Thus, the differences in average bequests across columns are relatively larger than the differences in average additional school years, in particular for low-ability quintiles.

The matrix shows that there is much less than full insurance against ability risk in the calibrated economy. For example, the cell in the first row and first column of the matrix shows that 86% of families currently in the lowest ability quintile are also in the lowest quintile of the welfare distribution. As shown in Table 3.4.1, the rank-rank correlation between ability and welfare within a generation is 0.90. More insurance across generations and less social mobility within a generation are two sides of the same coin.

We investigate further the mechanisms through which nurture affects the trans-

[19] We find that labor supply is approximately uncorrelated with ability in our calibrated model so that changes in welfare result mostly from changes in consumption.

mission of inequality. As mentioned above, bequests reduce labor effort through a wealth effect. This introduces an endogenous mean reversion in labor earnings because parents with higher labor earnings leave more bequests, thus inducing less labor effort and earnings of their children. We compute the average steady-state elasticity of labor earnings with respect to bequests to gauge how much the endogenous mean reversion reduces the persistence in labor earnings across generations. The bottom panel of Table 3.4.1 shows that the average elasticity of earnings with respect to bequests is indeed negative at -0.033. Interestingly the order of magnitude of this elasticity is in line with the evidence on lottery winners by Imbens et al. (2001) for the U.S. and by Cesarini et al. (2017) for Sweden. They estimate remarkably similar marginal propensities to earn out of changes in unearned income. Cesarini et al. (2017) report that (pre-tax) earnings decrease by 1.1 percent of the change in wealth, and this effect is very persistent. The average earnings response in our calibrated model is very similar at 1.5 percent.

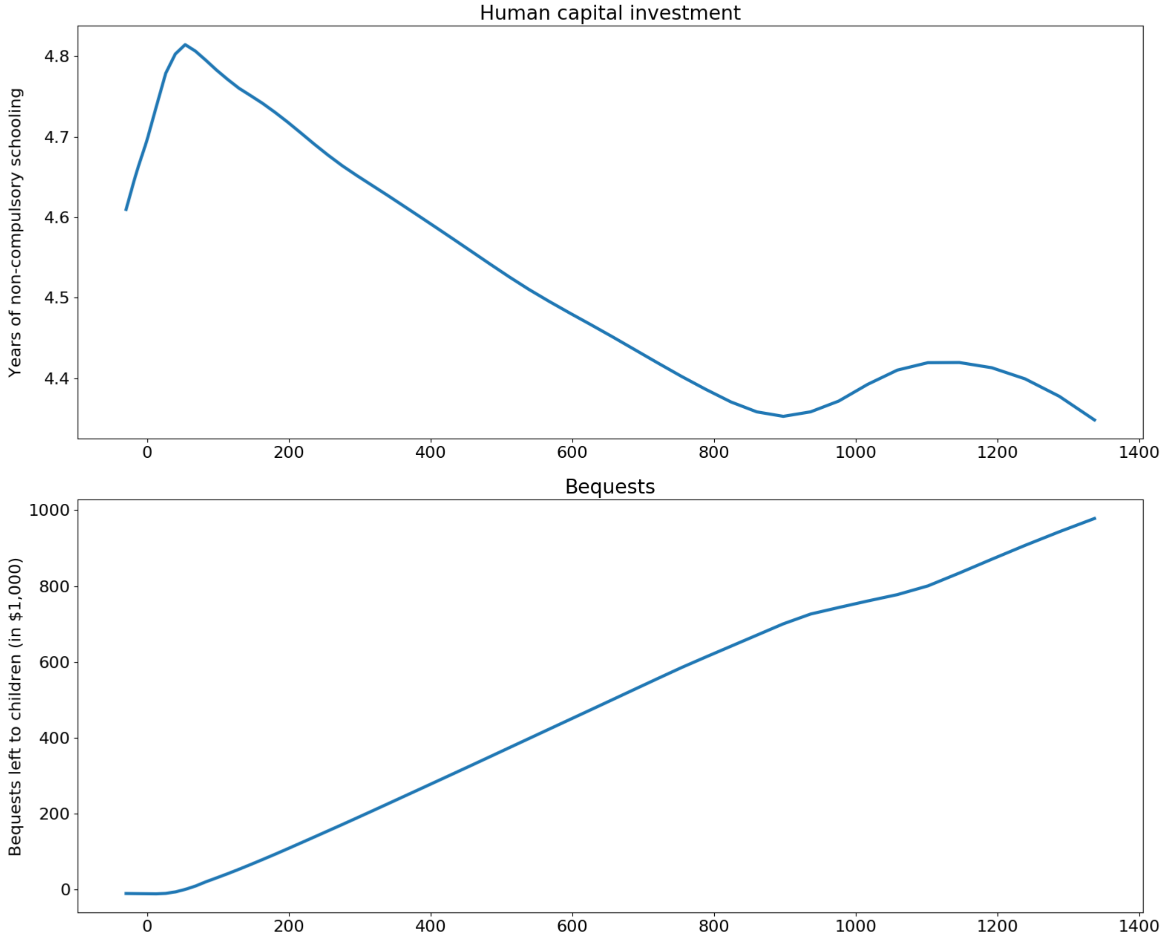


Figure 3.4.1: Human capital investment and bequests as a function of received bequests. *Notes:* We condition on parents with 12 years of schooling and median ability. Bequests are in units of $\$1,000$.

Another channel through which bequests affect mobility, is that bequests relax financial constraints for human capital investment. The relatively modest value of median bequests, resulting from a highly concentrated empirical distribution, implies that in the calibrated economy the spending of nearly half of the families is financially constrained.[20] To illustrate the consequences for human capital invest-

[20]The incidence of the financial constraint does not imply that the intergenerational earnings elasticity deviates much from the intergenerational correlation of ability. This is similar to Lee and Seshadri (2019) but for a different reason. As explained in Section 3.3, the empirical estimates on the returns to schooling imply that the effect of ability on labor productivity is much stronger than the effect of schooling on productivity in the calibrated model. In Appendix 3.9.5 we check robustness of our results if we target the mean of bequests, conditionally on receiving one. In this calibration 6.5% of families are financially constrained. As discussed at the end of Section 3.6, this increases the amount of insurance in the calibrated steady state relative to the benchmark cali-

ment, Figure 3.4.1 plots human capital investment and bequests for children as a function of bequests that parents received, for a representative family with median ability and high-school education. The figure shows that human capital investment is a very non-monotonic function of bequests because there are various effects at work. At a low level of bequests, human capital investment increases in bequests because the borrowing constraint is binding, illustrated by the flat portion of the policy function for the bequests left to children. More financial resources thus allow more human capital investment. Once the financial constraint is slack, the negative wealth effect on labor effort of the next generation implies that it is less attractive for parents to invest into their children's human capital: more schooling for the children only increases children's welfare in our economy if they work so that this investment generates income. Figure 3.4.1 further shows that once bequests start to be taxed, it becomes more attractive again to invest additional resources into schooling rather than to leave further bequests. This changes the slope of the plotted functions because relatively more human capital accumulation then ensures that the endogenous (risk-adjusted) return to human capital equals the after-tax return on bequests.

Compared to bequests, human capital affects the transmission of earnings inequality very differently in our calibrated economy. It makes labor earnings more persistent, less mobile and more unequal across generations. Given that parents with higher labor earnings invest more into the human capital of their children, high earnings are transmitted to their children as long as the substitution effect dominates so that an increase in labor productivity increases labor effort. The bottom panel of Table 3.4.1 shows that this is the case in the calibrated economy. The average elasticity of earnings with respect to schooling investment is 0.50, and we find that the elasticity has the highest value of 0.55 at the top of the earnings distribution where families have an ability above average. The elasticity of labor earnings with respect to nature (ability) is even higher at 0.93 because of the stronger effect of ability on labor productivity implied by the empirical estimates on the returns to schooling in the calibration.

---

bration. Qualitatively as in the benchmark case, social mobility decreases and insurance increases further on the transition path to the social optimum.

Given these transmission channels of nurture and nature, one may ask what size of changes in nurture and nature generates the same welfare effect. This is of interest from an intergenerational insurance perspective. To answer this question, we report results of the following experiment in Table 3.4.3. Consider a family characterized by the initial conditions (b, h, θ). Then compute the welfare increase if that family receives additional $\$10,000$, for example as bequest. Table 3.4.3 displays the increase in years of schooling or the increase in ability, in units of its standard deviation, which would generate the same welfare increase. In the different rows of the table, we show the average results of this experiment for families in different quintiles of the earnings distribution. We have chosen the earnings distribution because earnings are observable but the results are very similar for quintiles of the welfare distribution, as shown in Table 3.9.2 in Appendix 3.9.4.[21]

Earnings quintile	Increase in years of schooling	Cost of additional schooling	Increase in ability (in units of standard deviation)
1	0.97	19,393	0.08
2	0.70	13,772	0.09
3	0.57	11,429	0.09
4	0.47	9,489	0.10
5	0.34	7,106	0.11

Table 3.4.3: Average increase in schooling and ability that is welfare equivalent to receiving an additional $10,000 as bequests, by earnings quintile

The results in Table 3.4.3 illustrate the effectiveness of nurture (b, h) relative to nature θ in generating the same amount of welfare. In terms of efficiency, the increase in schooling, reported in the first column of the table, implies direct costs for the current generation that are reported in the second column.[22] The direct costs of the additional years of schooling, that are welfare-equivalent to obtaining additional $\$10,000$ as bequest, are larger than $\$10,000$ for families in the lower three quintiles of the earnings distribution. These families have relatively less ability, given the

[21]The correlation between earnings and welfare is very high at 0.81 so that the values in the last column of Tables 3.4.3 and 3.9.2 do not differ up until the second decimal place.

[22]For simplicity, we do not consider in these calculations that more human capital reduces the cost of investing into the human capital of the next generation. The reported costs can thus be considered an upper bound. Note that the cost of an additional year of schooling are approximately $\$20,000$ and thus higher than the cost of an additional school year at high-school graduation targeted in the calibration. The reason is that many families invest into schooling beyond high-school graduation and that the cost of schooling is convex.

nearly perfect correlation between ability and earnings of 0.99. Thus, schooling is less effective for these families in generating welfare than bequests. For families in the top of the earnings distribution instead, schooling is more efficient than bequests for generating additional welfare, given the complementarity of ability and human capital in generating labor productivity. For these families, the welfare-equivalent direct costs of the additional schooling are smaller than $\$10,000$. Finally, the last column of Table 3.4.3 shows that the required changes of ability, which are equivalent in welfare terms to additional bequests of $\$10,000$, are larger for higher quintiles of the earnings distribution. This indicates the decreasing returns in ability and shows the extent to which a given shock to ability has a stronger welfare impact at the bottom of the earnings distribution.

After illustrating the mechanisms through which nurture and nature affect the transmission of inequality in the calibrated model, we compute the quantitiative effects of nature and nurture on labor earnings y and the intergenerational transmission of earnings. We find that the effect of nurture on earnings through schooling and bequests is modest in the calibrated model. The variation in bequests and schooling given to a generation explains $0.9 - 1.2$ percent of the cross-sectional variance of that generation's earnings, depending on whether the covariance is split proportionately or equally across the determinants. The bequests and schooling *received by parents* explain at most 4.8 percent of the part of the variation of children's earnings that can be attributed to parents' nature and nurture. These results are consistent with the important role of nature in the transmission of earnings emphasized in recent empirical research by Bingley et al. (2018) based on a credible "Children of Twins" design using unique Danish data. They find that two thirds of the intergenerational earnings elasticity can be attributed to nature. Our model attributes an even larger role to ability but this has to be interpreted as an upper bound when assessing the effect of nature, given that ability at labor market entry in our model may also contain some nurture component.

3.5 Social optimum

To put inequality and mobility in the calibrated economy into perspective, we will compare them to their respective counterparts in the social optimum. We assume that asymmetric information constrains the insurance or redistribution provided by a utilitarian planner who discounts the future and weighs family dynasties equally. This implies non-degenerate inequality and mobility in the social optimum so that comparison between the calibrated economy and the social optimum is non-trivial.[23]

We focus on the planner problem with full commitment which provides an upper bound for the amount of insurance the planner can provide given the constraints. In such an environment, Farhi and Werning (2007) have analyzed allocations chosen by a utilitarian planner who discounts the future less than family dynasties and weighs dynasties equally. Denoting the planner's discount factor with ψ, they considered the case in which $\psi > \beta$. They showed that this assumption breaks the immiseration result of Atkeson and Lucas (1992) and implies a non-degenerate stationary distribution of consumption and welfare in the planner problem.

We refer to Appendix 3.9.1 for details about the planner problem and its solution. We emphasize two key equations, derived in the appendix, which show how the parameters ψ and β shape the solution of the planner problem. Let $\mathbb{E}$ denote the expectation operator, $u'(\cdot)$ the marginal utility of consumption, and let $c_t(\theta^t)$ denote consumption at time t as function of the sequence of abilities θ^t until time t, which is truthfully revealed by families to the planner in equilibrium. In the social optimum,

$$\mathbb{E}\left[\frac{1}{u'(c_t(\theta^t))}\right] = \frac{1/(1+r)}{\psi}\mathbb{E}\left[\frac{1}{u'(c_{t+1}(\theta^t))}\right] . \tag{3.3}$$

Stationarity of consumption thus requires $\psi = 1/(1+r)$, i.e., the planner's discount rate has to equal the real interest rate. Furthermore, we obtain an equation based on expectations conditional on the sequence θ^t:

$$\mathbb{E}\left[\frac{1}{u'(c_{t+1}(\theta^{t+1}))}\middle|\theta^t\right] = \frac{\beta}{\psi}\frac{1}{u'(c_t(\theta^t))} + \eta_t\left(1-\frac{\beta}{\psi}\right), \tag{3.4}$$

[23] In the first best with full insurance all inequality in consumption among ex-ante identical households would be eliminated.

where η is the multiplier attached to the constraint that captures the differences in the promises made by the planner compared to the promises which the planner would make if the same discounting as the family were applied to the welfare of future generations. See the recursive problem and its constraints in Appendix 3.9.1.

Since $\beta/\psi < 1$ and the multiplier $\eta > 0$, equation (3.4) shows that $1/u'(c(\theta^t))$ is mean-reverting. Because the planner cares more about providing equal opportunities for new generations in the future, shocks to ability are not fully passed on to future generations. The mean reversion implies a form of insurance against ability risk in the social optimum which contrasts the increase of inequality required for incentive provision. For $\beta = \psi$ instead, $1/u'(c(\theta^t))$ would follow a martingale, implying immiseration as in Atkeson and Lucas (1992).

As Farhi and Werning (2007), pp. 375-376, we focus on the case $\psi = 1/(1+r)$ in which the social optimum implies a stationary distribution for consumption. Our calibration of $\beta(1+r) < 1$ then disciplines the extent to which the planner cares more about providing opportunities for future generations than a dynasty itself. A stationary distribution in the social optimum also makes it sensible, in our view, to analyze a tax reform which implements a transition towards the socially-optimal steady state, starting from the status quo characterized by the economy that we calibrated to the U.S.

3.6 Transition to the social optimum

We analyze the transition to the social optimum, starting from the steady state of the calibrated economy. We implement this hypothetical reform by ensuring that the allocation for each dynasty has the same present discounted value of the expected net costs in the planner problem as in the calibrated economy. Although there is no redistribution across dynasties at the time of reform, the planner may redistribute towards future generations within a dynasty, given that the planner cares relatively more about welfare of future generations ($\psi > \beta$). The design of the reform ensures that the effects on insurance and mobility are not confounded by wealth effects. Appendices 3.9.2, 3.9.3 together with equation (3.25) and the subsequent discussion in Appendix 3.9.1, provide further details on how we implement the numerical solution and the reform.

We focus on results on the transition path rather than in the steady state of the social optimum because our interest is on the changes of mobility and inequality in the first decades after the reform. We also report key results for the new steady state after the reform as further benchmark. The steady state is approximated by the period 100 generations after the reform. This is conservative because we have found that in our experiments convergence happens much faster.

The evolution of key variables of interest after the reform is intuitive. Average consumption increases by 1.2% after the reform illustrating the efficiency gains of the social optimum compared to the calibrated economy. Expenditures for human capital and average labor earnings increase although average labor effort decreases.

The distributions in Figure 3.9.1, Appendix 3.9.4, show that the averages hide substantial heterogeneity. The figure plots distributions up to two generations after the reform and reveals that labor supply and human capital become much more dispersed after the reform as the planner decouples production of families from their consumption. Figures 3.9.2 and 3.9.3 show the evolution of the distributions for families, conditioning on the top and bottom quartile of the ability distribution in each generation. The different changes of the distribution of labor supply across ability types in generation 0, when the reform is implemented and assets and human capital are still given by pre-reform decisions, shows that social optimality requires an increase of labor supply of high ability types and more investment into the human capital of their children (visible in the distribution of non-compulsory school years plotted for generation 1). Labor supply of low ability types is reduced instead so that low-ability families obtain more of their welfare through enjoying more leisure. Moreover, some dynasties with currently high ability previously had low ability and vice versa, which affects the resources implied by promised utility (which we continue to call bequests after the reform in the figures). Thus, labor supply and human capital expenditures become more dispersed within the top quartile of the ability distribution, and the dispersion of labor supply increases also within the bottom quartile. We now provide further evidence on the larger dispersion within ability types on the transition path to the social optimum by analyzing the effects of the reform on mobility.

Beyond the effects on inequality, we find that income mobility (or persistence of labor earnings) remains quite stable on the transition to the social optimum. Two generations after the reform (corresponding to $30-60$ years), for example, the rank-rank correlation of earnings is 0.39 compared to 0.43 in the calibrated economy and the rank-rank correlation of consumption is 0.77 compared to 0.64. The similar persistence of earnings across generations, accompanied by the increase of the persistence of consumption, suggests that more intergenerational insurance against ability shocks is provided after the reform, reducing social mobility within generations.

	Quintiles				
Ability / Welfare	1	2	3	4	5
1	0.38	0.24	0.17	0.13	0.07
	(-264.19,5.0)	(-256.43,5.2)	(-222.59,5.4)	(-122.23,5.5)	(111.82,5.5)
2	0.25	0.23	0.21	0.18	0.13
	(-272.46,5.1)	(-265.87,5.3)	(-240.52,5.5)	(-158.71,5.6)	(64.26,5.7)
3	0.18	0.21	0.22	0.21	0.18
	(-276.69,5.1)	(-270.64,5.3)	(-247.81,5.5)	(-174.55,5.6)	(43.01,5.7)
4	0.13	0.18	0.22	0.23	0.24
	(-278.87,5.1)	(-274.97,5.3)	(-254.32,5.5)	(-187.46,5.6)	(31.15,5.8)
5	0.06	0.13	0.19	0.25	0.37
	(-285.55,5.1)	(-281.24,5.3)	(-262.42,5.5)	(-201.38,5.7)	(28.66,5.9)

Table 3.6.1: Social mobility matrix on the transition to the social optimum (in t=1, i.e., the second generation after the reform)

Notes: See Table 3.4.2.

This suggestive evidence is confirmed comparing the social mobility matrix two generations after the reform, presented in Table 3.6.1, with the mobility matrix of the calibrated economy in Table 3.4.2. The mobility matrix after the reform has less weight on the diagonal, implying more insurance and less mobility. The correlation between the rank in the ability distribution and the rank in the welfare distribution is indeed much smaller at 0.38, compared with 0.90 in the calibrated economy.[24] More intergenerational insurance on the transition towards the social optimum is achieved with a larger dispersion of wealth and human capital across families with different abilities,[25] and the correlation between wealth (or promises) and human

[24] Given that we simulate the economy for a sample of $500,000$ families, this difference is statistically significant.

[25] In the social optimum, families do not face borrowing constraints as in calibrated economy so

capital falls after the reform: in the second generation after the reform ($t = 1$), $cor(b, \log(h)) = 0.12$ compared with 0.44 in the calibrated economy.

To further compare the role of parental background in the calibrated economy and on the transition path to the social optimum, we regress the rank in the welfare distribution on the ranks in the distributions for bequests, human capital, and ability. Table 3.6.2 displays the results for the calibrated economy in column (1) and for the economy two generations after the reform in column (2). The linear specification explains most of the variation in welfare ranks, e.g., 92% in the calibrated economy according to the R^2 statistic. The regression coefficients in Table 3.6.2 show how moving up one decile in the distribution of bequests, human capital, or ability, respectively, changes the rank in the welfare distribution. For example, the coefficient of 0.77 for θ in column (1) implies that if ability were one decile higher in the ability distribution, then the family would move up 0.77 deciles in the welfare distribution.

Rank in distribution of...	*Welfare rank* (1) Calibrated economy	(2) 2nd generation after reform	(3) steady state after reform
b	0.29	0.55	0.99
S	0.14	0.68	0.36
θ	0.77	0.02	0.01
N	500,000	500,000	500,000
R^2	0.92	0.89	0.97

Table 3.6.2: Welfare rank regressions

*Notes:*The estimation results are obtained with an OLS-regression of welfare ranks on an intercept and the ranks in the distributions of bequests (b), schooling (S), and ability (θ). Column (1) uses simulated data of the calibrated economy. Columns (2) and (3) use simulated data of the socially optimal economy in the second generation after the reform and in the steady state, respectively.

Table 3.6.2 indicates that, in the calibrated economy, nature θ plays a more important role for a family's place in the welfare distribution than nurture b or S. The results in column (1) show that the rank in the ability distribution is approximately 2.5 to 5.5 times as important as bequests or schooling, respectively. In the

that wealth (or promises) can take more negative values.

second generation after the reform ($t = 1$), more intergenerational insurance against ability shocks increases the importance of nurture relative to nature dramatically: as shown in column (2), nurture is two orders of magnitude more important than nature for the position in the welfare distribution. The results in column (3) show that the same is true in the new steady state after the reform where, among the two variables capturing the effect of nurture, bequests (or promises in the terminology of the planner problem) become more important than schooling for the position in the welfare distribution.

Table 3.6.3 summarizes the analysis by showing how insurance and mobility evolve after the reform, starting from the steady state of the calibrated economy. The pass-through of an unexpected change in productivity to consumption falls after the reform, illustrating the increase of consumption insurance. Column 1 in Table 3.6.3 shows that the pass-through coefficient, obtained from a linear regression of log consumption on log productivity (which equals the log wage under perfect competition), decreases from 0.35 in the calibrated economy to 0.28 in the second generation after the reform and 0.31 in the steady state. That is, in the steady state after the reform 69 percent of ability shocks are insured compared to 65 percent in the calibrated economy.[26] Similarly, the rank-rank correlation between ability and welfare within generations decreases. Instead, mobility of income across and within generations changes very little. In particular, the high rank-rank correlation between ability and labor earnings within generations, that remains close to one after the reform, shows that the additional insurance after the reform can be decoupled from efficiently higher production of more able families. At the same time, the transition after the reform is associated with an increase in inequality of earnings, labor effort, consumption and bequests, as measured by the respective standard deviation reported in Table 3.9.3 in Appendix 3.9.4.

[26]Although the size of the pass-through coefficient in the calibrated intergenerational economy is not comparable directly to estimates obtained in a life-cycle context, a common finding in the literature is that US households are insured against a substantial part of persistent shocks. See, e.g., Heathcote et al. (2014) for an analysis of partial insurance in a model with consumption and endogenous labor supply.

	Pass-through coefficient	Rank-rank correlation of ability and welfare	Rank-rank correlation of ability and labor earnings	IGE
Calibrated economy	0.35	0.90	0.99	0.45
Social optimum				
2nd generation after reform	0.28	0.38	0.99	0.42
Steady state after reform	0.31	0.24	0.96	0.44

Table 3.6.3: Insurance and Mobility

Notes: The pass-through coefficient captures the effect of unexpected changes in productivity on consumption, obtained from a linear regression of log consumption on log productivity. IGE denotes the intergenerational earnings elasticity.

Table 3.9.4 in Appendix 3.9.5 shows that the results on insurance and mobility reported in Table 3.6.3 are quantitatively robust if we target a lower intergenerational earnings elasticity of 0.3 instead of 0.45, if we calibrate the model for a higher Frisch elasticity of 0.86 instead of 0.5, and if we recalibrate the economy with a higher complementarity between ability and schooling than in the benchmark. We do not report the results on income mobility because it does not change much after the reform for all cases but for the case in which we target the mean instead of the median of bequests in the calibration. In this case, the larger amount of bequests in the calibrated economy implies that there is an extended time period after the reform in which these bequests are run down at the same time as human capital increases. This has a positive effect on labor supply and is associated with an increase of the persistence of labor earnings from 0.45 to 0.62. As shown in column 2 of Table 3.9.4 in Appendix 3.9.5, a higher target level of bequests reduces social mobility, in terms of the correlation between ability and welfare, and increases insurance in the calibrated economy compared to the benchmark case. Qualitatively as in the benchmark case, social mobility decreases and insurance increases on the transition path to the steady state in the social optimum.

Given that we have constructed the reform without redistribution of resources across dynasties from an ex-ante perspective, our results on insurance and mobility show to which extent, at the time of the reform, dynasties are willing to trade less mobility within a generation for more intergenerational insurance. As is common in settings with risk sharing and insurance, generations with high ability may be better off ex post with less insurance but are bound by the commitment in our model environment. Without such commitment less insurance could be achieved if generations with high ability would have to be made indifferent to an outside option, which they would obtain if they reneged on the risk sharing arrangement.[27]

It is worth noting that compared to the transition to a social optimum with immiseration, there is more social mobility and a smaller increase in inequality to maintain incentives. As discussed, for example in Kocherlakota (2010), pp. 158-159, the inequality-increasing incentive effect is kept in check by the motive to provide opportunities for later generations if, as in our model, the planner discounts the future less than families and not at the same rate as in models that imply a social optimum with immiseration.

3.7 Simple tax and subsidy systems and the social optimum

Wedges, as defined in Appendix 3.9.6, capture the differences between the laissez faire and the socially optimal allocation based on comparison of the respective first-order conditions. Non-zero wedges imply that choices in the laissez faire need to be modified by taxes or subsidies to implement the social optimum. We provide simple history-independent approximations of these taxes and subsidies and discuss to which extent they allow to achieve the welfare gains of the reform.

We compare economies with simple tax and subsidy systems to the social optimum using two complementary approaches. We solve for the optimal linear taxes and subsidies that do not vary across generations after the reform, based on the problem (3.32) specified in Appendix 3.9.6 and modified to include linear, constant rates for taxes and subsidies. As an alternative, we approximate linear schedules

[27]Lack of the commitment of the government to stick to the implemented tax schedules also would impose a further constraint to achieve ex-ante credibility of the implemented tax schedule, in the sense that deviations have to be made sufficiently unattractive ex post. Farhi et al. (2012) show that such a constraint induces progressivity of the marginal tax on capital and can make the level of the capital tax positive. As shown by Findeisen and Sachs (2018) this result may not extend to human capital.

that may vary across generations analogous to the approximation of Farhi and Werning (2013) or Stantcheva (2017) in a life-cycle context. We thus set the tax or subsidy rates for income, bequests and schooling in each generation to their cross-sectional weighted averages in the second best, as further explained in Appendix 3.9.6.[28]

	Bequest	Schooling	Labor Income
Optimal linear taxes / subsidies, constant across generations	-0.36	-0.30	0.18
Approximated linear taxes / subsidies, varying across generations:			
2^{nd} generation after reform	-0.21	-0.45	0.31
Steady state after reform	-0.23	-0.50	0.41

Table 3.7.1: Simple linear taxes and subsidies

Notes: The taxes and subsidies varying across generations are cross-sectional averages derived in Appendix 3.9.6. A positive value implies a tax, a negative value implies a subsidy.

Table 3.7.1 displays the resulting linear taxes or subsidies. Bequests and schooling are subsidized while labor income is taxed. The qualitative features of this tax and subsidy system are intuitive. The planner has a lower discount rate than the dynasties and thus wants to provide further incentives so that dynasties shift resources to the future. The Pigouvian subsidy of bequests thus aligns the dynasties' incentives with the social optimum. Human capital accumulation is subsidized as well. As shown in Bovenberg and Jacobs (2005) and illustrated in Appendix 3.9.6, this is done also to offset the distortion of the schooling decision resulting from the taxation of labor income. The size of the subsidy further depends on the riskiness of human capital and possible distortions of the accumulation motive or incentives as emphasized by Stantcheva (2017) and Koeniger and Prat (2018).

[28] Solving for the optimal taxes and subsidies is numerically feasible if we restrict the tax schedules to be linear and constant in the post-reform period. We then perform a global search for the three optimal tax or subsidy rates and locally apply the Nelder-Meade optimization algorithm. A higher dimensionality of the space of parameters characterizing the tax and subsidy system quickly makes this procedure prohibitively costly in terms of computing time, e.g. if one attempts to solve for an optimal non-linear tax and subsidy system or if one allows for different taxes and subsidies across generations on the transition to the social optimum.

Table 3.7.1 shows that bequests and schooling are subsidized more compared to the calibrated economy, in which bequests are taxed only above the large exemption of \$756, 000 and schooling subsidies are included in the net cost for education. Thus, the level of the subsidy rates reported in Table 3.7.1 corresponds to the quantitative difference of the rates to the calibrated economy. The first row of Table 3.7.1 further shows that the optimal subsidy rate for schooling is six percentage points smaller than for bequests.

Concerning labor income taxation, the optimal linear tax rate reported in the first row of Table 3.7.1 is eleven percentage points smaller than the average marginal income tax in the calibrated economy. The bottom part of Table 3.7.1 shows that, for the approximated tax and subsidy rates which vary across generations but are not optimized, the absolute size of the taxes or subsidies increases slightly on the transition to the new steady state to provide insurance through nurture in terms of bequests and human capital at the same time as the inequality in labor earnings increases. Comparing the results in the bottom part with those in the top part of Table 3.7.1 reveals that the level of the approximated rates differs substantially from the optimal but constant rates. We now show that the latter imply a better approximation of the social optimum in our application.

We assess to which extent the linear tax and subsidies achieve the welfare gains of the second-best allocation relative to the laissez-faire allocation.[29] We compute the welfare gains at the time of the reform. By the design of the reform, we keep the present value of the expected net cost of the allocation per dynasty unchanged at the time of reform to focus on insurance and mobility by abstracting from redistribution across dynasties.

The first row of Table 3.7.2 shows the welfare gains for the social optimum, in which the planner is more patient than the dynasties ($\psi > \beta$). We report the welfare gains in percentage changes of consumption equivalents from the perspective of the planner, applying the discount factor ψ.[30] The welfare gains of 4.8% are larger than

[29]We compare the welfare gains relative to the laissez faire without distortionary taxes, starting from the steady state of the calibrated economy. In the laissez-faire economy we add a lump-sum tax so that each dynasty of type (b,h,θ) contributes the same present value of net taxes as in the calibrated economy and thus, by design, also as in the second best. This prevents a wealth effect from biasing the welfare comparison. Appendix 3.9.3 contains further details.

[30]If we evaluate the welfare gains from the perspective of the family, applying the discount factor

the gains between 1% and 3% reported in the life-cycle models of Farhi and Werning (2013) and Stantcheva (2017). By replicating their gains, we have verified that the difference comes from the larger initial cross-sectional heterogeneity in our setting, given that we start from the calibrated steady-state distribution, and from the intergenerational rather than life-cycle model implying different parameter values for the variance and the persistence of the shocks. The wedge between the discount factor of the planner and the dynasties instead cannot explain the difference, as we discuss further below. If compared to the calibrated economy, the welfare gains of the reform are a bit higher at 5.7% because of the borrowing constraints that are present in the calibrated but not in the laissez-faire economy.

	Welfare gain in percent of consumption equivalents
Second best	4.8
Optimal linear taxes / subsidies, constant across generations	2.2
Approximated linear taxes / subsidies, varying across generations	1.5

Table 3.7.2: Welfare gains

Notes: Welfare gains compared to the laissez faire, holding constant the present value of expected net costs for each dynasty at the time of the reform. The consumption equivalents are computed holding labor supply constant and applying the discount factor ψ.

The second and third row of Table 3.7.2 show the welfare gains achieved in the economies with the simple linear tax and subsidy systems. The economy with the approximated, varying tax and subsidy rates, in the third row of Table 3.7.2, achieves 31% of the welfare gains obtained by moving from the laissez faire to the second best after the reform. The economy with optimized but constant linear taxes and subsidies, in the second row of Table 3.7.2, achieves 46% of the welfare gains instead. This is a sizable part but less than in the calibrated life-cycle models of

β in the objective function, the welfare gains are smaller by a factor 0.31 as gains accruing in the future then receive less weight.

Farhi and Werning (2013) and Stantcheva (2017) who find that simple linear taxes deliver more than 90% of the welfare gains.[31]

In order to gauge the importance of $\psi > \beta$ for these results, i.e., the difference of the discount factor of the dynasties β and the discount factor of the planner ψ, we also compute the welfare gains for the case $\psi = \beta$, keeping the interest rate r unchanged.[32] As mentioned before, $\psi = \beta$ implies immiseration in the social optimum. In this case, the welfare gains of the socially optimal allocation after the reform compared to the laissez faire are 6.7%. Approximated linear taxes and subsidies, which vary across generations, imply welfare gains of 1.2%, thus achieving 18% of the welfare gains from the laissez faire to the second best after the reform. These results are similar in terms of orders of magnitude to those reported in Table 3.7.2 and illustrate that our conclusions for the benchmark case do not depend on the wedge between the discount factor of the planner and the dynasties in a critical way.

Although a substantial part of the welfare gains can be generated with the simple taxes and subsidies, one may expect that the history dependence of optimal taxes and subsidies is not fully captured by the endogenous state variables bequests b and human capital h in our model because the unobserved shocks to ability θ are persistent in our model and not i.i.d. as in Albanesi and Sleet (2006). Allowing for further history dependence of the taxes and subsidies, while maintaining tractability, seems viable within a life cycle of a generation, as shown in Kapička (2017), but less so across generations where this would require information on past generations for determining taxes and subsidies of the current generation.

[31] We have found that an approximated quadratic tax schedule, which captures the progressive phasing out of bequest subsidies emphasized by Farhi and Werning (2010), does not achieve higher welfare gains. This may be a consequence of approximation error and, unfortunately, computing the optimal quadratic tax schedule is prohibitively costly. Appendix 3.9.6 provides further details on the non-linear approximation.

[32] For a stationary distribution in the calibrated economy, we need to maintain the assumption that $\beta < 1/(1+r)$.

3.8 Conclusion

We have analyzed inequality and mobility in a dynasty model that we calibrated to match empirical evidence for the U.S. We have compared mobility and the intergenerational transmission of inequality in the steady state of the calibrated economy and on the transition to the social optimum, based on a social welfare function in which dynasties are weighed equally and, as a result, future generations receive a larger weight than in the welfare function of dynasties. The wedge between the weight of future generations in the dynasties' problem and the planner's problem has been disciplined by observable data on bequests in our calibration.

We have found that, compared with the calibrated U.S. economy, social optimality implies less opportunity within a generation, measured in terms of less mobility of agents with different ability across the welfare distribution, and more intergenerational insurance against ability risk. This is achieved with a stronger influence of nurture, in terms of schooling and bequests, on the family's position in the welfare distribution. We find that income mobility across and within generations remains quite stable, indicating that consumption can be decoupled from labor earnings in the social optimum.

We show that simple linear schedules for taxes and subsidies achieve about half of the welfare gains of the socially optimal tax system compared to the laissez faire. On average, bequests and schooling are subsidized in the social optimum to provide insurance of future generations against ability risk. The optimal linear rates of the simple tax and subsidy system are 18% for the labor income tax, 36% for the bequest subsidy and 30% for the schooling subsidy.

Our analysis of the transition to the social optimum shows that less social mobility and more inequality of labor earnings ex post do not imply necessarily lower welfare ex ante because they may be the flip side of more intergenerational insurance. This illustrates that interpretation of descriptive evidence on the evolution of inequality and mobility require the usual assumptions about preferences and technology, and about the social welfare function. For one plausible set of these assumptions, we have shown how intergenerational insurance and mobility may be shaped by the tax and subsidy system.

3.9 Appendix

3.9.1 Planner problem: recursive formulation and results

In this appendix we present the problem of a utilitarian planner who maximizes the welfare of generations under incentive compatibility constraints. The socially-optimal solution of this problem requires that families truthfully reveal their hidden ability. We first present both the primal and the dual problem of the planner. We then provide the recursive formulation of the relaxed dual problem, based on the first-order approach.[33] We use the first-order conditions of the relaxed problem to derive the key equations (3.3) and (3.4) discussed in Section 3.5 of the main text. Stating the problem of the planner requires that we discuss incentive compatibility.

Incentive compatibility.—We focus on a direct revelation mechanism which ensures that families truthfully report their type in each generation. We denote the history of types within a given family as $\theta^t \equiv \{\theta_0, \theta_1, ..., \theta_t\}$ and history dependent allocations as $x_t(\theta^t) \equiv \{c_t(\theta^t), h_{t+1}(\theta^t), y_t(\theta^t)\}$. The feasible set $\mathcal{X}$ contains all sequences $\mathbf{x} \equiv \{x_t(\theta^t)\}_{t=1}^T$ of measurable functions $x_t : \Theta^t \to \mathbb{R}_+^3$. Using the production function to substitute l_t in the utility function and writing $U(c_t, y_t, h_t, \theta_t)$ instead of $\mathbf{U}(c_t, l_t)$, preferences of a family dynasty for an allocation $\mathbf{x}$ are

$$\mathcal{U}(\mathbf{x}) \equiv \mathbb{E}_0\left[\sum_{t=1}^{\infty} \beta^{t-1} U\left(c_t\left(\theta^t\right), y_t\left(\theta^t\right), h_t\left(\theta^{t-1}\right), \theta_t\right)\right],$$

where $\mathbb{E}_0$ is the expectation operator conditional on information available at time 0 and β is the discount factor of the family.

Family dynasties can choose any reporting strategy $\mathbf{r} \equiv \{r_t(\theta^t)\}_{t=1}^T$ from the set $\mathcal{R}$ containing all sequences of measurable functions $r_t : \Theta^t \to \Theta$. The types are private information so that an allocation must be incentive compatible to ensure truthful reporting, i.e.,

$$(IC) : \mathcal{U}(\mathbf{x}) \geq \mathcal{U}(\mathbf{x} \circ \mathbf{r}), \text{ for all } \mathbf{r} \in \mathcal{R}, \tag{3.5}$$

[33] This approach replaces the incentive-compatibility constraints with an envelope condition that needs to be satisfied on the equilibrium path on which families truthfully reveal their type.

where $(\mathbf{x} \circ \mathbf{r})(\theta^t) \equiv \{x_t(r^t(\theta^t))\}_{t=1}^{\infty}$ is the allocation $\mathbf{x}$ resulting from the reporting strategy $\mathbf{r}$ and history θ^t.

Primal problem.—We assume a utilitarian planner who weighs the welfare of each family dynasty equally and discounts the future less than a dynasty, i.e. $\psi > \beta$. See Kocherlakota (2010), chapter 5, p. 146, for a textbook treatment. The planner can fully diversify the idiosyncratic ability risk that family dynasties face. Since there are no general equilibrium feedbacks that link the problems of the dynasties, the planner can maximize aggregate welfare by maximizing welfare of each dynasty. For a utilitarian planner, the problem of insuring family dynasties under the veil of ignorance (from the perspective of period 0) is equivalent to the problem of optimal redistribution across family dynasties with different initial conditions. The primal problem of the planner is

$$W = \max_{\{c_t, y_t, h_{t+1}\}} \mathbb{E}_0 \sum_{t=1}^{\infty} \psi^{t-1} U(c_t, y_t, h_t, \theta_t) \tag{3.6}$$

$$\text{s.t. } \mathbb{E}_0 \sum_{t=1}^{\infty} \beta^{t-1} U(\cdot) \geq V, \tag{3.7}$$

$$(IC),$$

$$\mathbb{E}_0 \sum_{t=1}^{\infty} q^{t-1} z_t \leq \Gamma_0,$$

where z_t is the per-period net cost (i.e., $z_t \equiv c_t + g(h_{t+1}, h_t) - y_t$), Γ_0 is a given level of average discounted costs, V is a (promised) utility level and $q \equiv 1/(1+r)$. Without loss of generality, we can assume an initial (distribution of) promised utility.

We consider the constant discount factor ψ in the planner's objective. In the derivation of the planner's objective in Kocherlakota (2010), p. 147, the discount factor is time varying and converges to ψ if $\psi > \beta$ and $t \to \infty$. See also Kocherlakota (2010), p. 157. We abstract from possible time variation in the planner's discount factor assuming that the planning objective has converged. This has the advantage that our transition analysis after the tax reform in Section 3.6 is not confounded by changes in the discount factor over time on the transition path.

Dual problem.—The dual cost minimization problem of the planner is

$$\Gamma_0 = \min_{\{c_t, y_t, h_{t+1}\}} \mathbb{E}_0 \sum_{t=1}^{\infty} q^{t-1} z_t \tag{3.8}$$
$$\text{s.t. } \mathbb{E}_0 \sum_{t=1}^{\infty} \beta^{t-1} U(\cdot) \geq V$$
$$(IC),$$
$$\mathbb{E}_0 \sum_{t=1}^{\infty} \psi^{t-1} U(\cdot) \geq W.$$

Incentive constraints for the recursive formulation and the first-order approach.—We replace the ex-ante incentive constraint (3.5) with an ex-post requirement to write the planner's dual problem in recursive form.[34] For this purpose, we define the equilibrium continuation utility $\omega(\theta^t)$ for a given history θ^t as

$$\omega\left(\theta^t\right) \equiv U\left(c_t\left(\theta^t\right), y_t\left(\theta^t\right), h_t\left(\theta^{t-1}\right), \theta_t\right) + \beta \int_{\Theta} \omega\left(\theta^t, \theta_{t+1}\right) dF\left(\theta_{t+1} \middle| \theta_t\right), \tag{3.9}$$

for all $t = 1, ..., \infty$. Families compare the continuation value $\omega(\theta^t)$ of truthful reporting to the values derived from arbitrary reporting strategies

$$\omega^{\mathbf{r}}\left(\theta^t\right) \equiv U\left(c_t\left(r^t\left(\theta^t\right)\right), y_t\left(r^t\left(\theta^t\right)\right), h_t\left(r^{t-1}\left(\theta^{t-1}\right)\right), \theta_t\right) + \beta \int_{\Theta} \omega^{\mathbf{r}}\left(\theta^t, \theta_{t+1}\right) dF\left(\theta_{t+1} \middle| \theta_t\right).$$

Incentives are compatible ex-post if

$$\omega\left(\theta^t\right) \geq \omega^{\mathbf{r}}\left(\theta^t\right), \text{ for all } \theta^t \text{ and all } \mathbf{r} \in \mathcal{R}. \tag{3.10}$$

We use $\mathbf{x}^{\mathbf{IC}}$ to denote the set of all allocations $\mathbf{x}$ satisfying (3.10).[35]

Problem (3.8) requires to keep track of all the out-of-equilibrium payoffs to check the incentive constraint (3.10). Applying the first-order approach, we reduce the

[34]In this part we draw heavily on material in Koeniger and Prat (2018) which we present here for completeness.

[35]Note that allocations in $\mathbf{x}^{\mathbf{IC}}$ are incentive compatible for all $\theta^t \in \Theta^t$. This requires truth telling to be optimal after *any history* of shocks, whereas the incentive constraint (3.5) only requires truth telling to be ex-ante optimal. The two notions can only differ on a set of measure zero histories. In other words, allocations that are ex-ante incentive compatible are also ex-post incentive compatible almost everywhere.

complexity of the problem by replacing the incentive constraint with an envelope condition which only depends on the marginal utility of truthtellers. As in Koeniger and Prat (2018), the envelope condition for the considered problem is

$$\frac{\partial \omega\left(\theta^{t}\right)}{\partial \theta_{t}}=\frac{\partial U\left(c_{t}\left(\theta^{t}\right), y_{t}\left(\theta^{t}\right), h_{t}\left(\theta^{t-1}\right), \theta_{t}\right)}{\partial \theta_{t}}+\beta \int_{\Theta} \omega\left(\theta^{t+1}\right) \frac{\partial f\left(\theta_{t+1} \mid \theta_{t}\right)}{\partial \theta_{t}} d \theta_{t+1}. \tag{3.11}$$

Intuitively, if one considers a one-shot perturbation of the type θ_t in equation (3.9), the sum of all the derivatives of terms with respect to the report of the type is zero, once the derivatives are evaluated on the equilibrium path where truthful reporting is optimal. Equation (3.11) reduces to the condition prevailing in Mirrlees' static setting if types are i.i.d. In this case, the second term on the right-hand side of (3.11) vanishes. The second term on the right-hand side is relevant instead if types are persistent because unobserved ability then generates additional private information. For example, parents who underreport their type become more optimistic than the planner about the ability of their children if types are positively correlated.

Replacing the incentive constraint by (3.11) greatly simplifies the optimization problem because it only depends on the continuation utility of truthtellers and not on the continuation utility of all possible types. Defining $\mathbf{x}^{\mathbf{FOA}}$ as the set of allocations so that condition (3.11) holds for all θ^t, we note that $\mathbf{x}^{\mathbf{IC}} \subseteq \mathbf{x}^{\mathbf{FOA}}$. Replacing the incentive constraint in problem (3.8) by $\mathbf{x} \in \mathbf{x}^{\mathbf{FOA}}$ thus relaxes this problem.

Recursive relaxed problem.—We write the relaxed problem in recursive form so that we can solve it as sequence of standard optimal control problems. Denoting with "$'$" values of variables one period in the future and with "$-$" values of variables with a one-period lag, the stationary recursive problem is:

$$\Gamma\left(V, W, \Phi, h, \theta_{-}\right)=\min_{\{c, y, h', V', W', \Phi'\}}\Big\{$$
$$\int_{\Theta}\left[c(\theta)+g(h'(\theta), h)-y(\theta)+q \Gamma\left(V'(\theta), W'(\theta), \Phi'(\theta), h'(\theta), \theta\right)\right] d F(\theta | \theta_{-})\Big\},$$

$$\begin{aligned}
\text{s.t.} \qquad \omega(\theta) &= U(c(\theta), y(\theta), h, \theta) + \beta V'(\theta), &(3.12)\\
\tilde{\omega}(\theta) &= U(c(\theta), y(\theta), h, \theta) + \psi W'(\theta) &(3.13)\\
V &= \int_{\Theta} \omega(\theta) dF(\theta|\theta_{-}), &(3.14)\\
W &= \int_{\Theta} \tilde{\omega}(\theta) dF(\theta|\theta_{-}), &(3.15)\\
\Phi &= \int_{\Theta} \omega(\theta) \frac{\partial f(\theta|\theta_{-})}{\partial \theta_{-}} d\theta, &(3.16)\\
\frac{\partial \omega(\theta)}{\partial \theta} &= \frac{\partial U\left(c, y, h, \theta\right)}{\partial \theta} + \beta \Phi'(\theta). &(3.17)
\end{aligned}$$

Equations (3.12) and (3.13) define the continuation values using the discount factor of the family and planner, respectively. Equations (3.14) and (3.15) are the respective promise keeping constraints. Because of the persistence of ability, the planner keeps track how reports of ability in the last period change promised utility so that the problem also has a threat-keeping constraint (3.16). The envelope condition (3.17) is the incentive compatibility constraint as explained above.

The recursive problem is standard but for the additional constraints (3.13) and (3.15) which enter the problem because the planner discounts the future at a different rate than the family dynasties. Substituting (3.13) into (3.15), and substituting $U(\cdot)$ using (3.12) and (3.14), we obtain

$$W = V + \int_{\Theta} \big(\psi W'(\theta) - \beta V'(\theta)\big) dF(\theta|\theta_{-}). \tag{3.18}$$

The problem has therefore one more state variable W and equation (3.18) as additional constraint, which replaces (3.13) and (3.15). The additional state variable and constraint (3.18) would be redundant if the planner and the dynasties discounted the future at the same rate because $W = V$ and $\int_{\Theta} \psi W'(\theta) - \beta V'(\theta) dF(\theta|\theta_{-}) = 0$ in this case.

Optimality conditions.—We use the separability of utility in consumption and labor effort to solve constraint (3.12) for consumption. We then substitute the resulting consumption $c(\omega(\theta) - \beta V'(\theta), y(\theta), h, \theta)$ into the objective function. The

Hamiltonian reads

$$\begin{aligned}\mathcal{H} =&[c(\omega(\theta) - \beta V'(\theta), y(\theta), h, \theta) + g(h'(\theta), h) - y(\theta) \\ &+ q\Gamma(V'(\theta), W'(\theta), \Phi'(\theta), h'(\theta), \theta)] f(\theta, \theta_-) \\ &+ \lambda[V - \omega(\theta) f(\theta, \theta_-)] + \gamma \left[\Phi - \omega(\theta) \frac{\partial f(\theta, \theta_-)}{\partial \theta_-}\right] \\ &+ \eta \left[W - V - \big(\psi W'(\theta) - \beta V'(\theta)\big) f(\theta|\theta_-)\right] \\ &+ \mu(\theta) \left[\frac{\partial U(c(\omega(\theta) - \beta V'(\theta), y(\theta), h, \theta), y(\theta), h, \theta)}{\partial \theta} + \beta \Phi'(\theta)\right] .\end{aligned} \tag{3.19}$$

The first-order conditions for h', y and Φ' remain qualitatively unchanged compared with those reported in Appendix A of Koeniger and Prat (2018). We thus focus on the first-order necessary conditions which generate new insights. The first-order condition for V' is

$$\frac{\beta}{\partial u(c(\theta))/\partial c(\theta)} - \beta\eta = q \frac{\partial \Gamma(V'(\theta), W'(\theta), \Phi'(\theta), h'(\theta), \theta)}{\partial V'(\theta)} \tag{3.20}$$

and the first-order condition for W' is

$$\eta\psi = q \frac{\partial \Gamma(V'(\theta), W'(\theta), \Phi'(\theta), h'(\theta), \theta)}{\partial W'(\theta)}. \tag{3.21}$$

We now use these two equations to derive a modified reciprocal Euler equation and the key equations (3.3) and (3.4) discussed in Section 3.5 of the main text. As a first step, we substitute the envelope condition $\partial\Gamma/\partial W = \eta$ into equation (3.21) which implies

$$\eta\psi = q\eta'(\theta). \tag{3.22}$$

The shadow price of the new constraint (3.18) thus evolves deterministically which simplifies the numerical solution.

Next, we substitute the envelope condition $\partial\Gamma/\partial V = \lambda - \eta$ into first-order condition (3.20) to obtain

$$\frac{1}{\partial u(c(\theta))/\partial c(\theta)} = \frac{q}{\beta}\lambda'(\theta) - \frac{q}{\beta}\eta'(\theta) + \eta. \tag{3.23}$$

Using (3.22) to substitute $\eta'(\theta)$,

$$\frac{1}{\partial u(c(\theta))/\partial c(\theta)} = \frac{q}{\beta}\lambda'(\theta) + \eta\left(1 - \frac{\psi}{\beta}\right). \tag{3.24}$$

This is the modified reciprocal Euler equation because $\lambda'(\theta) = \mathbb{E}[(\partial u(c'(\theta'))/\partial c'(\theta'))^{-1}|\hat{\theta}]$, as can be shown following the same steps as in the proof of Remark 1 in Koeniger and Prat (2018), for example. Note that the expectation conditions on the history of abilities until the current period which we denote by $\hat{\theta}$.

As a further step, we take unconditional expectations in equation (3.24) and rearrange to obtain

$$\mathbb{E}\left[\frac{1}{\partial u(c'(\theta'))/\partial c'(\theta')}\right] = \frac{\beta}{q}\frac{1}{\partial u(c(\theta))/\partial c(\theta)} + \eta\left(\frac{\psi}{q} - \frac{\beta}{q}\right). \tag{3.25}$$

For the implementation of the reform, as discussed in further detail in Section 3.9.3, we set $\eta = \lambda$ at the time of the reform. We now show that this choice implies that average consumption changes at the time of the reform and remains constant afterwards at its steady state level if $\psi = q$, i.e., if the planner's discount rate equals the real interest rate r. Equation (3.25) implies that for $\eta = \lambda$ and $\psi = q$,

$$\mathbb{E}[(\partial u(c'(\theta'))/\partial c'(\theta'))^{-1}] = \mathbb{E}[\lambda'] = \eta. \tag{3.26}$$

Iterating forward, using (3.25) with a one-period lead, $\eta = \mathbb{E}[\lambda']$ and $\psi = q$, it follows that average consumption remains constant after an initial adjustment at the time of the reform.

Long-run properties.— If we take unconditional expectations on both sides of (3.24) and rearrange using that $\mathbb{E}[\eta] = \mathbb{E}[\lambda]$, we obtain key equation (3.3) in Section

3.5 of the main text, denoting the marginal utility of consumption as $u'(\cdot)$:

$$\mathbb{E}\left[\frac{1}{u'(c_t(\theta^t))}\right] = \frac{q}{\psi}\mathbb{E}\left[\frac{1}{u'(c_{t+1}(\theta^{t+1}))}\right]$$

Clearly, stationarity of consumption requires $\psi = q$. Substituting this into the modified reciprocal Euler equation (3.24), we obtain key equation (3.4) in Section 3.5 of the main text:

$$\mathbb{E}\left[\frac{1}{u'(c_{t+1}(\theta^{t+1})}\middle|\theta^t\right] = \frac{\beta}{\psi}\frac{1}{u'(c_t(\theta^t))} + \eta_t\left(1 - \frac{\beta}{\psi}\right).$$

For $\beta/\psi < 1$, $1/u'(c(\theta^t))$ follows a mean-reverting process. This recovers, for our model setting, results in Farhi and Werning (2007), p. 380, or Kocherlakota (2010), p. 158.

Note that we get the standard reciprocal Euler equation if $\eta = 0$, in which case the constraint for providing a certain amount of welfare W is slack, or if $\psi = \beta$ so that the immiseration result applies. As we have just seen, stationarity of consumption in the planner problem requires $\psi = q$; and a stationary distribution in the decentralized calibrated economy with incomplete markets requires $q > \beta$. Thus, $\psi > \beta$ seems a rather natural assumption.

The steady state consumption level is determined by the resources of the planner, as mentioned in Farhi and Werning (2007), p. 385, and $\eta_0 = \partial\Gamma_0/\partial W_0$ measures the marginal cost for the planner of providing social welfare W_0 in period 0. We can index the planner problem by η_0 since the entire sequence of multipliers η_t is deterministic. This property not only simplifies the numerical solution but also the implementation of the reform, as we explain further in Appendix 3.9.3.

3.9.2 Numerical solution

We solve problem of dynasties applying the endogenous gridpoint method, as explained in the online Appendix C of Koeniger and Prat (2018). The planner problem is solved building on the programs of Farhi and Werning (2013).

For the dynastic problem we use a grid of 75 points for bequests, 100 points for human capital, and 12 points for ability θ. Consistent with our interpretations of

$\log(h)$ as non-compulsary schooling years in the Mincer wage regression, the lowest grid point of human capital is $\exp(0) = 1$. Note that the cost function thus implies a very small minimum expenditure for education over a 30-year period of \$600 for the calibrated parameter values.

We calibrate the problem of dynasties by minimizing $D = \sum_{i=1}^{k} \left(\frac{x_i^m - x_i^d}{x_i^d}\right)^2$ where x_i^m is the i-th moment generated by the model and x_i^d is the corresponding target moment in the data. We compute a moment x_i^m in the stationary distribution by simulating the model for 500,000 dynasties. To calibrate the model, we start by performing a global search on a parameter grid to minimize this expression. From the best parameters thus obtained we then use the Nelder-Meade optimization algorithm to further improve on the fit of the model. With this strategy we are able to reduce D close to zero, i.e. to fit the data precisely.

For the numerical solution of the planner problem, we follow Farhi and Werning (2013), pp. 614-615, and replace the state variables (V, Φ) with the multipliers (λ, γ). This has computational advantages because the domain of the multipliers is bounded below. Furthermore, conditioning the problem on λ allows us to solve the first-order conditions to determine the allocation and then obtain V by integrating once over the utility of that allocation. Similarly, conditional on γ, we can obtain Φ by integrating once. This speeds up the numerical solution because otherwise we would have to ensure, for example, that the chosen allocation integrates to V and computationally expensive integration potentially would have to be performed many times.

We choose a grid of 17 points for λ, 12 points for γ, 18 points for h, 25 points for θ_{-}, and 26 points for ϵ. All programs are implemented in *julia*. On a standard processor of the current vintage, solving the calibrated economy takes 2 minutes and solving the planner problem takes 30 hours.

3.9.3 Implementation of the reform to attain the social optimum

In this section we describe how we construct the reform to attain the social optimum starting from the steady state of the calibrated economy. We first simulate the calibrated economy for $M = 500,000$ households. We approximate the stationary distribution by simulating the economy for 100 generations. We label the steady-state of the calibrated economy as $t = -1$. The planner proposes the reform at $t = 0$. Note that the reform is proposed after consumption and human capital investment decisions by the parents have been made but before their childrens' types are realized and therefore before the children make their decisions.

To focus on the implications of the reform for insurance, we abstract from redistribution between dynasties by holding constant the present discounted value of net costs of each dynasty's allocation. The reported welfare gains are thus not confounded by wealth effects. We now describe in more detail how we implement the reform.

Resource constraint.— In the calibrated economy, households pay a positive amount of net taxes to the government. These taxes can be thought of being used to finance an exogenous stream of government expenditures which in the calibrated economy amount to 27 percent of average labor earnings. Since we do not model the expenditure side of the government, we assume that the planner has to continue to raise the amount of resources required for these government expenditures. In other words, the net government surplus in the reformed economy equals the surplus in the calibrated economy so that the planner does not have more resources available due to some arbitrary assumption about a change in the size of that surplus in the reform. Each dynasty in state $s = (b, h, \theta)$ at the time of the reform has to contribute the same amount of net-taxes $t(s)$ as in the calibrated economy. Then the planner's allocation satisfies for each state

$$\widehat{\Gamma}_\eta(\lambda(s), 0, h(s), \theta) = (1 + r)b(s) + t(s), \tag{3.27}$$

where $\widehat{\Gamma}_\eta(\cdot)$ is the expected cost for the planner of providing the allocations for the dynasties conditional on the state variables of the relaxed planner problem derived in Appendix 3.9.1. Note that $\widehat{\Gamma}_\eta(\cdot)$ is the cost function for any value of η once the state variables (V, Φ) have been replaced by their multipliers (λ, γ), analogous to Farhi and Werning (2013). The multiplier $\gamma(s) = 0$ for all dynasties in state s because, apart from the utility promise and the parents' ability, there is no further restriction from history so that the threat-keeping constraint is not binding in the reform period.

The remaining degree of freedom in equation (3.27) is η_0 at the time of the reform which depends on the resources available to the planner. We set $\eta_0(s) = \lambda_0(s)$ which, by the envelope condition $\partial\Gamma/\partial V = \lambda - \eta$ obtained in Appendix 3.9.1, implies that a marginal variation of the promised utility V leaves the planner's cost for the allocation of a dynasty in state s unchanged.

Assets.— Given that the socially optimal allocation does not determine dynasties' assets or bequests, we briefly mention how we compute them after the reform. If one interprets assets as the difference between the net present value of expenditures and the net present value of earnings, as frequently done in the literature, the counterpart of assets in the planner's problem is the expected present value of net costs $\tilde{\Gamma}(s)$ for providing an allocation.

3.9.4 Further results

y_t / y_{t+1}	Quintiles 1	2	3	4	5
1	0.42	0.25	0.17	0.11	0.05
2	0.24	0.25	0.22	0.18	0.11
3	0.16	0.21	0.22	0.22	0.18
4	0.11	0.17	0.21	0.25	0.27
5	0.07	0.12	0.17	0.24	0.40

Table 3.9.1: Earnings quintile transition matrix in the steady state of the calibrated economy

Notes: The probabilities across columns in each row may not add up to 1 because of rounding.

Welfare quintile	Increase in years of schooling	Cost of additional schooling	Increase in ability (in standard deviations)
1	0.94	17,465	0.08
2	0.68	12,977	0.09
3	0.57	11,317	0.09
4	0.48	10,251	0.10
5	0.40	9,177	0.11

Table 3.9.2: Average increase in schooling and ability that is welfare equivalent to receiving an additional $10,000 as bequests, by welfare quintile

	Standard deviations of logarithms				
	Bequests	Schooling	Labor effort	Earnings	Consumption
Calibrated economy	0.16	0.08	0.06	0.46	0.41
Social optimum					
2$^{\text{nd}}$ generation after reform	0.39	0.06	0.10	0.51	0.36
Steady state of reform	1.04	0.07	0.16	0.53	0.55

Table 3.9.3: Standard deviations in the calibrated and reformed economy

Notes: For bequests we report the standard deviation of the level and not the logarithm because the variable bequests can take negative values. Bequests after the reform correspond to the present value of net costs for an allocation, as explained further in Appendix 3.9.3.

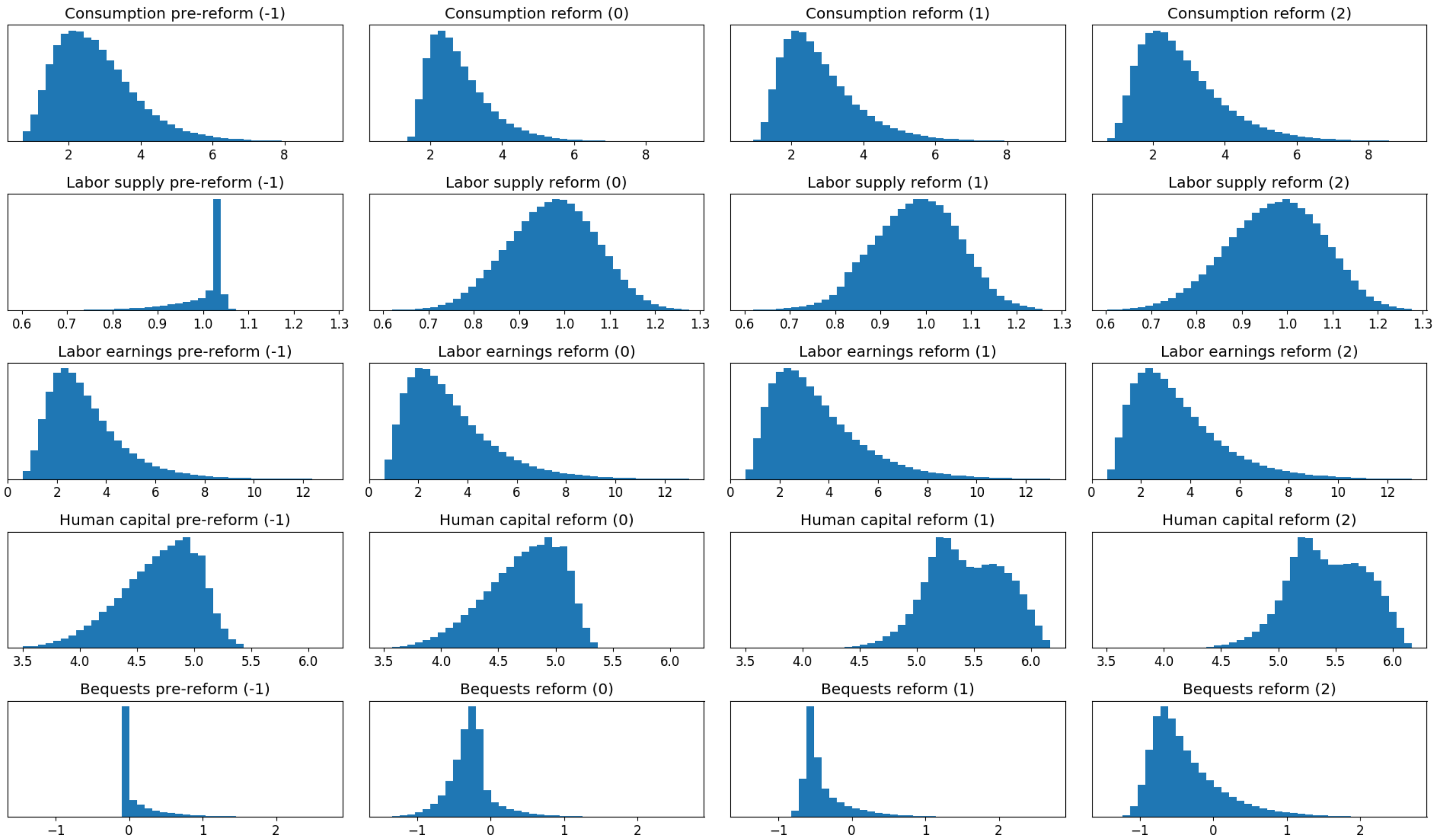


Figure 3.9.1: Transition to the social optimum: evolution of distributions. *Notes:* The reform is implemented before generation 0 makes its choices. Units of monetary variables are in $10,000$ US-\$ and labor supply is normalized by the average level in the calibrated steady state.

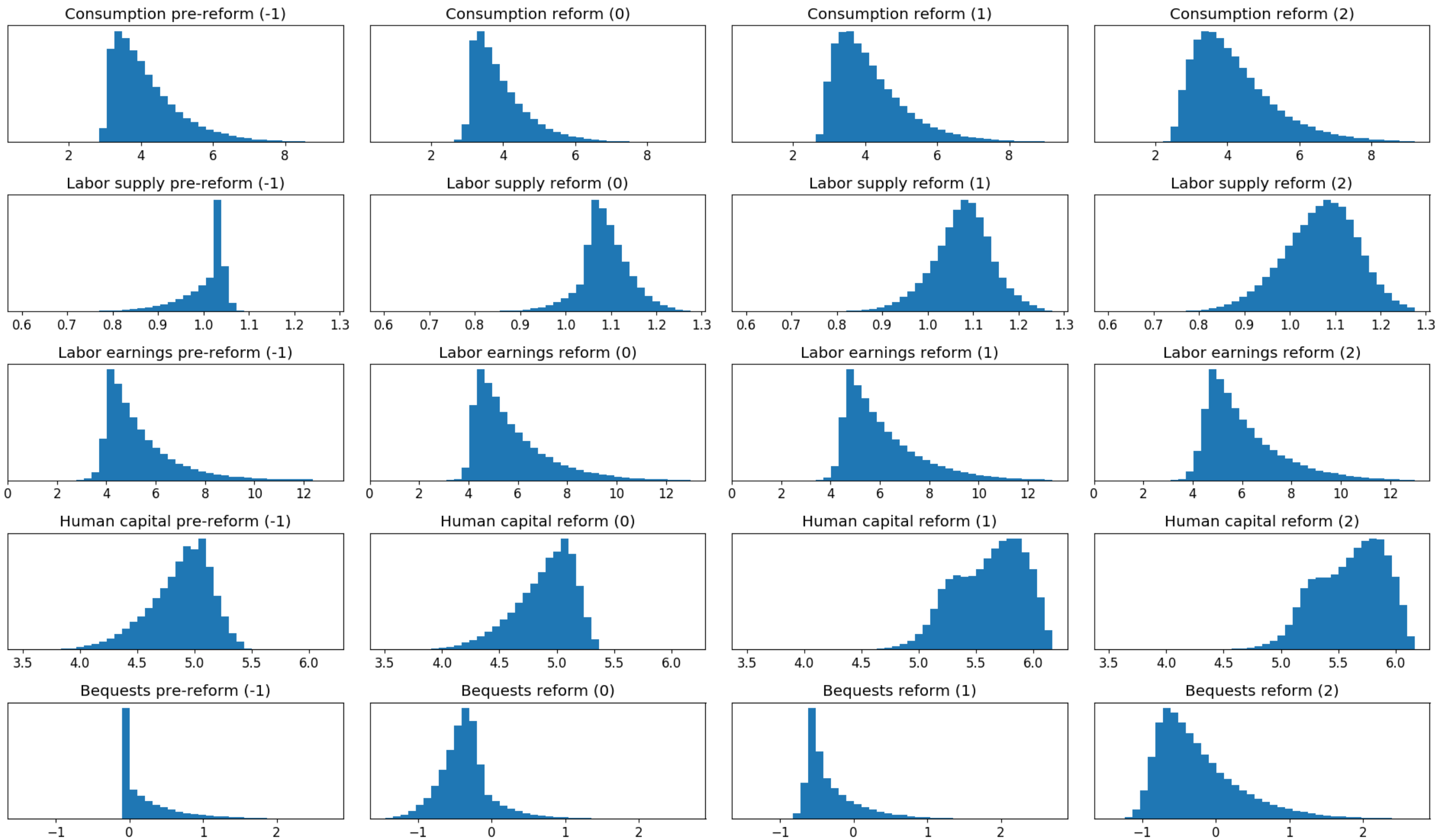


Figure 3.9.2: Transition to the social optimum: evolution of distributions for top quartile of ability distribution in each generation. *Notes:* See Figure 3.9.1.

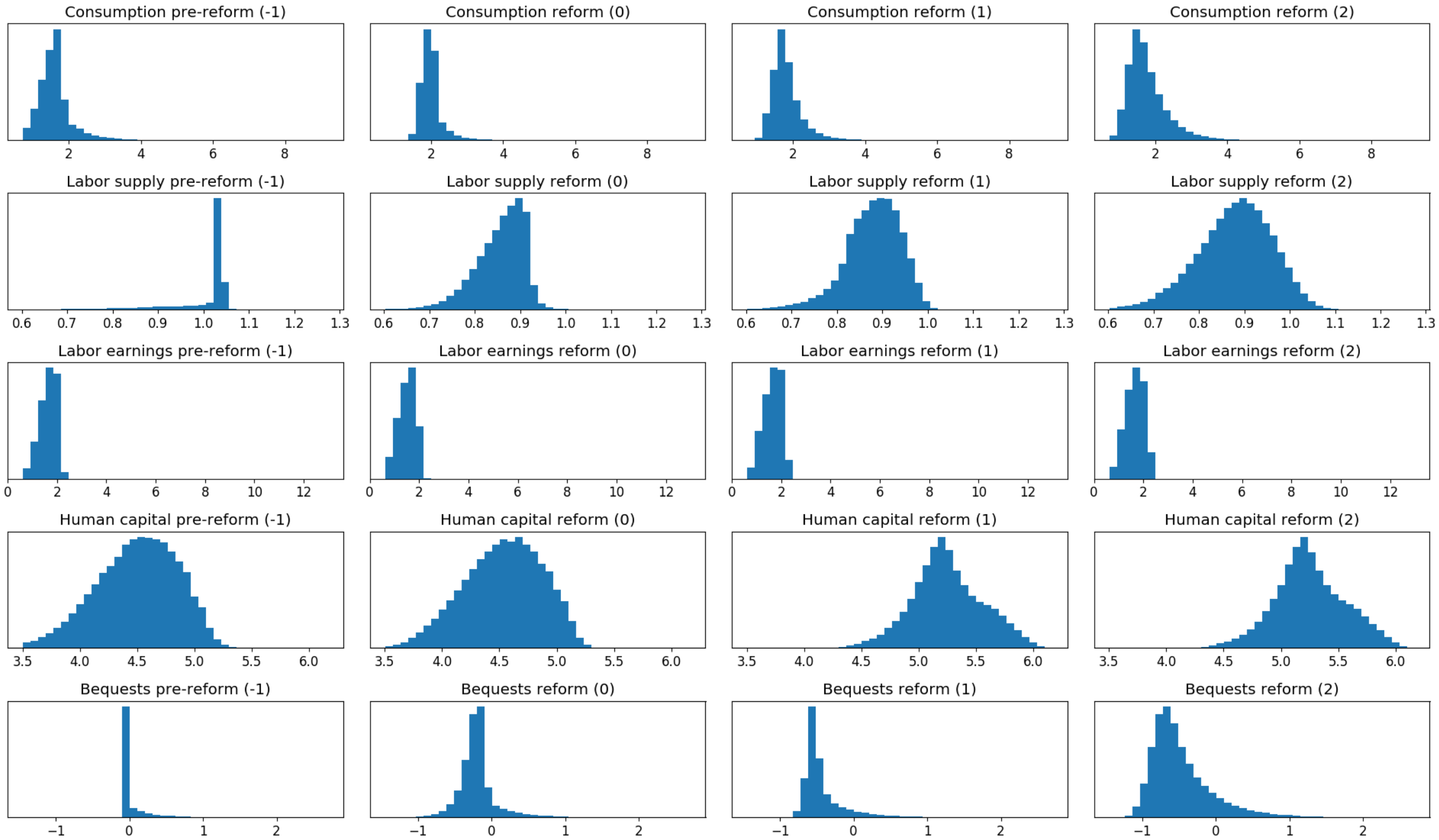


Figure 3.9.3: Transition to the social optimum: evolution of distributions for bottom quartile of ability distribution in each generation. *Notes:* See Figure 3.9.1.

3.9.5 Robustness analysis

This subsection presents the results of the robustness analysis. Table 3.9.4 shows that the results on the evolution of insurance and mobility for our benchmark calibration, presented in Table 3.6.3 in the main text, are robust across most alternative calibrations. For convenience, we repeat the results of the benchmark case in the first column of Table 3.9.4. Worth noting is that social mobility, in terms of the correlation between ability and welfare, decreases and insurance increases in the calibrated economy if we target the conditional mean of bequests in column (2). This calibration implies more bequests than in the benchmark calibration and thus less dependance of welfare on current ability.

Table 3.9.5 provides the corresponding details on the results of the calibration for each of the considered robustness checks, in terms of the recalibrated parameter values and the implied target statistics. We now provide further information for each of the robustness checks.

Mean bequests as target.— We target the mean bequests of households that received a bequest. As before, we convert the mean bequest of $408,400 for households, reported in table 2 of Wolff and Gittleman (2014), into adult equivalents dividing by 1.4 so that our target is $291,714. As shown in the second column of Table 3.9.5, the calibration matches this target quite closely.

Lower intergenerational elasticity of earnings as target.— We recalibrate the model if we target a lower intergenerational earnings elasticity $\iota = 0.3$, which is at the low end of estimates reported in Chetty et al. (2014), table 1. The third

	(1)	(2)	(3)	(4)	(5)
	Benchmark	*Mean bequests as target*	*Lower IGE as target*	*Higher Frisch elasticity*	*Higher complementarity between θ and h*
Pass-through coefficient					
Calibrated economy	0.35	0.19	0.34	0.37	0.27
Social optimum					
2nd generation after reform	0.28	0.18	0.26	0.32	0.21
Steady state after reform	0.31	0.30	0.29	0.36	0.22
Rank-rank corr(ability,welfare)					
Calibrated economy	0.90	0.49	0.87	0.92	0.95
Social optimum					
2nd generation after reform	0.38	0.19	0.26	0.42	0.38
Steady state after reform	0.24	0.12	0.17	0.26	0.27

Table 3.9.4: Robustness of results for insurance and social mobility

	(1) Benchmark	(2) Mean bequests as target	(3) Lower IGE as target	(4) Higher Frisch elasticity	(5) Higher complementarity between θ and h
Recalibrated parameters					
Discount factor (annualized) β	0.966	0.970	0.965	0.965	0.963
Persistence ρ	0.45	0.44	0.33	0.49	0.44
Education cost parameter κ	0.0014	0.0015	0.0015	0.0017	0.0017
Education cost parameter ς_1	0.75	0.78	0.78	0.68	0.70
Education cost parameter ς_2	-0.0005	-0.0489	-0.0514	0.0443	0.0177
Predictions for target statistics					
Bequests	52,772	298,458	52,681	51,678	52,711
Average years of schooling S	12.75	12.73	12.77	12.71	12.88
Correlation(S', S)	0.48	0.47	0.44	0.44	0.46
Intergenerational earnings elasticity	0.45	0.43	0.30	0.48	0.44
Average net cost of an additional year of schooling	13,954	13,691	13,696	13,821	13,725

Table 3.9.5: Calibration results for robustness checks

column of Table 3.9.5 shows that the recalibrated model continues to match the target statistics closely. Furthermore, the model-implied mobility matrix, displayed in Table 3.9.6, matches more closely the estimated transition matrix reported in Table 2 of Chetty et al. (2014). The match of the rank-rank correlation of income also improves, as mentioned in footnote 18.

Higher Frisch elasticity.— We recalibrate the model for a larger Frisch elasticity of 0.86 for aggregate hours, reported in Chetty et al. (2013), table 2. This elasticity is based on micro-estimates from quasi-experimental studies and contains responses of hours at the intensive and extensive margin. We thus set $\alpha = 2.16$. The fourth column of Table 3.9.5 shows that in the recalibrated model persistence of the ability shock slightly increases to match the intergenerational earnings elasticity. This quantitative result obtains because labor supply is not only a function of ability but also of bequests and human capital. If labor supply instead were a power function of ability, then one can show that a higher Frisch elasticity would only affect the variance of log income but not the persistence of log income across generations. Further note that ς_2 is calibrated to be positive, implying that families with lower human capital have a cost advantage for educating their children (which can be motivated, for example, because of lower opportunity costs).

Higher complementarity between human capital and ability.— We recalibrate our model to match the complementarity between years of schooling and ability, as suggested by findings in the second row of table 3 in Cunha et al. (2006). Using test results of the AFQT as a measure of ability, they find that the return to one year of college in percent at the 95-th percentile of the ability distribution is 1.6 times higher than the return at the 5-th percentile of the ability distribution (and not constant across the ability distribution as implied by the Cobb-Douglas assumption for productivity). We use this target to calibrate the elasticity of substitution χ between ability θ and human capital h in the function for productivity. Because wages are no longer log-separable in ability and years of schooling if productivity is not Cobb-Douglas, we also calibrate the variance of the innovation of the ability process and the parameter ξ using the variance of residual wages of 0.2, obtained from a Mincer wage regression on the years of schooling, and the return to schooling of 0.1 as target statistics.[36] The calibration of these parameters is done jointly with the other parameters reported in Table 3.9.5. Our recalibration results in $\chi = 0.786$, $\sigma_\epsilon^2 = 0.295$ and $\xi = 0.73$ compared with $\chi = 1$, $\sigma_\epsilon^2 = 0.217$ and $\xi = 0.9$ in the benchmark calibration. The additional targets are matched well by the recalibrated model: the return to the first year of college at the 95-th percentile of the ability distribution is 1.6 times higher than the return at the 5-th percentile of the ability distribution, the variance of residual wages is 0.2 (both statistics equal the respective target up to three digits of precision), and the return to schooling is 0.09. The fifth column of Table 3.9.5 shows that the recalibrated model also continues to match the other targets reasonably well where ς_2 is calibrated to be positive, as in the robustness check with the higher Frisch elasticity.

[36]Given that we observe ability in our simulated data based on the model, we obtain the model counterpart for the unbiased empirical estimate of the return to schooling by controlling for ability in the regression. In the empirical literature, identification of the return to schooling is based on a instrumental-variable regression given that ability is not fully observable and correlated with schooling.

	Quintiles				
y_t / y_{t+1}	1	2	3	4	5
1	0.35	0.24	0.19	0.14	0.08
2	0.22	0.23	0.22	0.19	0.14
3	0.17	0.20	0.21	0.22	0.20
4	0.14	0.18	0.20	0.23	0.25
5	0.12	0.15	0.18	0.23	0.32

Table 3.9.6: Earnings quintile transition matrix, in the steady state of the calibrated economy, targeting an intergenerational elasticity of earnings of 0.3

Notes: The probabilities across columns in each row may not add up to 1 because of rounding.

3.9.6 Details on the wedges and the approximate implementation

In the laissez faire each dynasty solves the maximization problem[37]

$$\begin{aligned}
\widetilde{W}\left(b,h,\theta_{-}\right) &= \max_{\{b'(\theta),h'(\theta),l(\theta),c(\theta)\}} \left\{ \int_{\Theta} \left[\mathbf{U}\left(c(\theta),l(\theta)\right) + \beta \widetilde{W}\left(b'(\theta),h'(\theta),\theta\right)\right] dF\left(\theta \middle| \theta_{-}\right) \right\} \\
\text{s.t. } b'(\theta) &= (1+r)b - c(\theta) - g(h'(\theta),h) + y(\theta), \\
y(\theta) &= Y(h,\theta,l(\theta)), \\
\ln(\theta) &= \rho \ln(\theta_{-}) + \epsilon,
\end{aligned}$$

where the family chooses functions $b', h', l, c : \Theta \to \mathbb{R}$. Note that, as in the planner problem but differently to the calibrated economy presented in Section 3.2, we make the common assumption that the dynasty faces no borrowing constraint in the laissez faire. Thus, below we obtain the standard definitions of the wedges based on the first-order conditions of the laissez faire problem.

[37]The first-order conditions of this problem are equivalent for the problem $W^L\left(b(\theta),h(\theta),\theta\right) = \max_{\{b'(\theta),h'(\theta),l(\theta),c(\theta)\}} \left\{ \left[\mathbf{U}\left(c(\theta),l(\theta)\right) + \beta \int_{\Theta} W^L\left(b'(\theta),h'(\theta),\theta'\right)\right] dF\left(\theta'|\theta\right)\right\}$, s.t. the constraints.

The first-order conditions for bequests, human capital and labor supply are:

$$\begin{aligned}
\frac{\partial \mathbf{U}(c,l)}{\partial c} &= \beta(1+r)\mathbb{E}\left[\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right], \\
\frac{\partial g(h',h)}{\partial h'}\frac{\partial \mathbf{U}(c,l)}{\partial c} &= \beta\mathbb{E}\left[\left(\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right)\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right], \\
-\frac{\partial \mathbf{U}(c,l)}{\partial l} &= \frac{\partial y}{\partial l}\frac{\partial \mathbf{U}(c,l)}{\partial c}.
\end{aligned}$$

Based on these first-order conditions and given the separability of the utility function, the history-dependent wedges at time t for bequests $\tau_{b,t}$, human capital $\tau_{h,t}$ and labor supply $\tau_{l,t}$ are then defined as

$$\tau_{b,t}\left(\theta^t\right) \equiv 1 - \frac{q}{\beta}\frac{\partial u\left(c_t\left(\theta^t\right)\right)/\partial c_t}{\mathbb{E}\left[\partial u\left(c_{t+1}\left(\theta^{t+1}\right)\right)/\partial c_{t+1}\middle|\theta^t\right]}, \tag{3.28}$$

$$\tau_{h,t}\left(\theta^t\right) \equiv \frac{\beta}{\frac{\partial g(h_{t+1}(\theta^t),h_t(\theta^{t-1}))}{\partial h_{t+1}}} \tag{3.29}$$

$$\cdot\,\mathbb{E}\left[\frac{\frac{\partial u\left(c_{t+1}\left(\theta^{t+1}\right)\right)}{\partial c_{t+1}}}{\frac{\partial u(c_t(\theta^t))}{\partial c_t}}\left(\frac{\partial y_{t+1}\left(\theta^{t+1}\right)}{\partial h_{t+1}} - \frac{\partial g(h_{t+2}\left(\theta^{t+1}\right),h_{t+1}\left(\theta^t\right))}{\partial h_{t+1}}\right)\middle|\theta^t\right] - 1, \tag{3.30}$$

$$\tau_{l,t}\left(\theta^t\right) \equiv 1 - \frac{\partial v(y_t\left(\theta^t\right),h_t\left(\theta^{t-1}\right),\theta_t)/\partial y_t}{\partial u\left(c_t\left(\theta^t\right)\right)/\partial c_t}, \tag{3.31}$$

where the function $v(\cdot)$ denotes the disutility of labor once we have substituted labor using the production function.

The social optimum can be decentralized with a general, history-dependent tax schedule as shown, for example, in Stantcheva (2017). We will approximate the implementation of the social optimum with a history-independent and linear tax schedule. With this goal in mind, we now specify an auxiliary, decentralized problem for each dynasty that helps us to explain how we approximate the linear history-independent tax schedule. The auxiliary problem is

$$\max_{\{b_{t+1},h_{t+1},l_t,c_t\}} \mathbb{E}_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \mathbf{U}(c_s, l_s) \right] \tag{3.32}$$

$$\begin{aligned} \text{s.t. } b_{t+1} &= (1+r)\left(1 - t_{b,t}\left(\theta^{t-1}\right)\right) b_t + \left(1 - t_{y,t}\left(\theta^t\right)\right) y_t \qquad (3.33) \\ &\quad - c_t - \left(1 + t_{h,t}\left(\theta^t\right)\right) g(h_{t+1}, h_t) - T\left(\theta^t\right), \\ y_t &= Y(h_t, \theta_t, l_t), \\ \ln(\theta) &= \rho \ln(\theta_-) + \epsilon, \end{aligned}$$

where the tax shifter $T(\theta^t)$ becomes a transfer if negative, and the conditioning of that shifter and the taxes $t_{j,t}(\cdot)$, $j = b, h, y$, on the history imply general, possibly non-linear tax schedules across dynasties with different histories. These general tax schedules allow to implement the social optimum and this can be achieved also by conditioning taxes on the history of observable variables such as output and education expenditures. As discussed in Stantcheva (2017), for example, this requires that the history of these observable variables allows to identify θ^t.

Note that $t_{b,t}(\theta^{t-1})$ is the tax rate applied at the time parents choose bequest b_t and $t_{h,t}(\theta^t)$ increases the cost of a year of schooling, $g(h_{t+1}, h_t)$, so that $t_{h,t}(\theta^t) < 0$ has the interpretation of a subsidy for human capital investment h_{t+1}. For realism, we implement the conditioning of taxes or transfers on human capital by linking the tax or subsidy for human capital to education expenditures.

We proceed by linking the taxes $t_{j,t}(\cdot)$, $j = b, h, y$ to the respective wedges and then use these relationships to approximate linear, history-independent taxes.

Labor income tax.—The first-order condition for labor supply or income y and the definition of the labor wedge (3.31) imply that the marginal income tax of a dynasty with a certain history θ^t equals the labor wedge, i.e. $t_{y,t}(\theta^t) = \tau_{l,t}(\theta^t)$. In Section 3.7 we are interested in how well the social optimum can be approximated with simpler linear taxes that do not depend on history. Given the non-tractability of solving for the optimal linear taxes, as mentioned in the main text, we proceed as Farhi and Werning (2013) or Stantcheva (2017) and approximate the linear income taxes with the cross-sectional average of the labor wedge:

$$\bar{t}_{y,t} = \mathbb{E}\left[\tau_{l,t}\left(\theta^t\right)\right] \tag{3.34}$$

so that every dynasty faces the same labor income tax.

Bequest tax.— The first-order condition with respect to b_{t+1} implies that

$$\frac{\partial u\left(c_t\right)}{\partial c_t} = \beta(1+r)\left(1 - t_{b,t+1}\left(\theta^t\right)\right)\mathbb{E}_t\left[\frac{\partial u\left(c_{t+1}\right)}{\partial c_{t+1}}\right]. \tag{3.35}$$

Thus, comparison with equation (3.28) implies that bequests of a dynasty with a certain history θ^t should be taxed at rate $\tau_{b,t}(\theta^t)$.[38] We approximate the bequest tax with the cross-sectional average of the bequest wedge, i.e.,

$$\bar{t}_{b,t+1} = \mathbb{E}\left[\tau_{b,t}\left(\theta^t\right)\right]. \tag{3.36}$$

Human capital tax.— We combine the first-order condition for human capital

$$\begin{aligned}\frac{\partial g(h_{t+1},h_t)}{\partial h_{t+1}}\left(1 + t_{h,t}\left(\theta^t\right)\right)\frac{\partial u\left(c_t\right)}{\partial c_t} =& \beta\mathbb{E}_t\left[\left(\frac{\partial y_{t+1}}{\partial h_{t+1}}\left(1 - t_{y,t+1}\left(\theta^{t+1}\right)\right)\right.\right. \\ &\left.\left. - \frac{\partial g(h_{t+2},h_{t+1})}{\partial h_{t+1}}\left(1 + t_{h,t+1}\left(\theta^{t+1}\right)\right)\right)\frac{\partial u\left(c_{t+1}\right)}{\partial c_{t+1}}\right]\end{aligned} \tag{3.37}$$

with the definition of the wedge for human capital (3.30) and solve for $t_{h,t}\left(\theta^t\right)$:

$$t_{h,t}\left(\theta^t\right) = \tau_{h,t}\left(\theta^t\right) - \frac{\beta}{\frac{\partial g(h_{t+1},h_t)}{\partial h_{t+1}}} \tag{3.38}$$

$$\cdot\,\mathbb{E}_t\left[\left(\frac{\partial y_{t+1}}{\partial h_{t+1}}t_{y,t+1}\left(\theta^{t+1}\right) + \frac{\partial g(h_{t+2},h_{t+1})}{\partial h_{t+1}}t_{h,t+1}\left(\theta^{t+1}\right)\right)\frac{\frac{\partial u(c_{t+1})}{\partial c_{t+1}}}{\frac{\partial u(c_t)}{\partial c_t}}\right]. \tag{3.39}$$

The equation shows that a positive wedge for human capital does not necessarily

[38] As discussed in Kocherlakota (2010), taxation of assets generally has to be implemented *ex post*, after realization of θ_{t+1}, thus ensuring that the Euler equation of families is satisfied for each consumption level at the reported ability. We approximate this *ex-post* heterogeneity in the tax rate when we consider non-linear taxes below.

imply a positive current marginal tax on human capital accumulation. The second term inside the expectation operator on the right-hand side shows that the sign and size of the tax also depends on how human capital changes labor income and thus labor-income taxes in the next period, how human capital changes the cost for education in the next period, and how these changes are correlated with the marginal utility of consumption. In particular, the planner has to undo the distortion on human-capital accumulation implied by labor-income taxation, as shown in Bovenberg and Jacobs (2005), and the distortion implied by the tax/subsidy on human capital next period. Note that through the stochastic discount factor, any distortion of the bequest decision also influences the tax/subsidy on human capital. Furthermore, effects of human capital accumulation on incentives are captured as well through the wedge $\tau_{h,t}$. Such effects occur if productivity is not Cobb-Douglas as emphasized, for example, by Stantcheva (2017).

As for the other taxes, we approximate the linear tax or subsidy for human capital investment by taking the cross-sectional average, i.e.,

$$\bar{t}_{h,t} = \mathbb{E}\left[t_{h,t}\left(\theta^t\right)\right]. \tag{3.40}$$

Given the recursive nature of equation (3.39), we use the approximated taxes in $t+1$ when approximating taxes in t. This ensures consistent use of the linear tax approximations in the dynasties' problem to compare the decentralized problem with the approximated linear tax schedule to the social optimum. We thus solve problem (3.32) replacing the general tax schedules with the approximated linear taxes $\bar{t}_{j,t},\ j = b, h, y$, and compare the welfare gains of this economy with simple linear taxes to the welfare gains of the social optimum with implicit non-linear and history-dependent taxes.

For the welfare comparisons of the decentralized economy with a simple non-linear approximation of the tax schedules, mentioned in footnote 31, we assume $t_{y,t}\left(y_t, \theta^t\right)$, $t_{b,t+1}\left(b_{t+1}, \theta^t\right)$ and $t_{h,t}\left(h_{t+1}, \theta^t\right)$ in the auxiliary problem (3.32). I.e., we capture explicitly some of the non-linearities of the tax schedules by letting the income tax depend on the current income, by letting the bequest tax depend on the bequest level and by letting the subsidies for human capital expenditures depend on the size of these expenditures. This allows, for example, to capture explicitly the progressivity emphasized by Farhi and Werning (2010) in the context of bequest taxation. The tax schedules still condition on the history given that the ability shocks are persistent and not i.i.d. so that the history is not fully encoded in the endogenous state variables. The first-order conditions of the auxiliary problem (3.32), with the modified tax schedules, and the definition of the wedges then imply that

$$t_{y,t}\left(y_t,\theta^t\right)+\frac{\partial t_{y,t}\left(y_t,\theta^t\right)}{\partial y_t}y_t=\tau_{l,t}\left(\theta^t\right), \tag{3.41}$$

$$t_{b,t}\left(b_t,\theta^t\right)+\frac{\partial t_{b,t}\left(b_t,\theta^t\right)}{\partial b_t}b_t=1-\frac{q}{\beta}\frac{\partial u\left(c_{t-1}\left(\theta^{t-1}\right)\right)/\partial c_{t-1}}{\partial u\left(c_t\left(\theta^t\right)\right)/\partial c_t}, \tag{3.42}$$

and

$$\begin{aligned} t_{h,t}\left(h_{t+1},\theta^t\right)+\frac{\partial t_{h,t}\left(h_{t+1},\theta^t\right)}{\partial h_{t+1}}\frac{1}{\epsilon_{g(\cdot),h_{t+1}}}h_{t+1}&=\tau_{h,t}\left(\theta^t\right) \\ -\frac{\beta}{\frac{\partial g(h_{t+1},h_t)}{\partial h_{t+1}}}\mathbb{E}_t\Bigg[\Bigg(\frac{\partial y_{t+1}}{\partial h_{t+1}}\Bigg(t_{y,t+1}\left(y_{t+1},\theta^{t+1}\right)&+\frac{\partial t_{y,t+1}\left(y_{t+1},\theta^{t+1}\right)}{\partial y_{t+1}}y_{t+1}\Bigg) \\ +\frac{\partial g(h_{t+2},h_{t+1})}{\partial h_{t+1}}t_{h,t+1}\left(h_{t+2},\theta^{t+1}\right)\Bigg)\frac{\frac{\partial u(c_{t+1})}{\partial c_{t+1}}}{\frac{\partial u(c_t)}{\partial c_t}}\Bigg]&, \end{aligned} \tag{3.43}$$

where $\epsilon_{g(\cdot),h_{t+1}}\equiv\frac{\partial g(h_{t+1},h_t)}{\partial h_{t+1}}\frac{h_{t+1}}{g(h_{t+1},h_t)}$ is the elasticity of education expenditures with respect to human capital. Note the right-hand side of (3.42) differs from the right-hand side of (3.28) because we approximate the non-linear tax schedule for bequests based on the implementation discussed in Kocherlakota (2010), which conditions on the current realization of the shock. See also Farhi and Werning (2010), p. 664, for this type of implementation in an intergenerational model.

Approximating the non-linearity of each tax with a quadratic function, the respective tax *rate* is linear, i.e., $t(x)=\widetilde{\alpha}+\widetilde{\beta}x$. The left-hand side of equations (3.41), (3.42) and (3.43) then becomes

$$\widetilde{\alpha}_{y,t}+2\widetilde{\beta}_{y,t}y_t, \tag{3.44}$$

$$\widetilde{\alpha}_{b,t} + 2\widetilde{\beta}_{b,t} b_{t+1}, \tag{3.45}$$

and

$$\widetilde{\alpha}_{S,t} + \widetilde{\beta}_{S,t} \frac{1}{\varsigma_1} + \widetilde{\beta}_{S,t} S_{t+1}, \tag{3.46}$$

where (3.46) follows from the parametric assumption for the cost function $g(h', h) = \kappa(h')^{\varsigma_1} h^{\varsigma_2}$ and from the change of variable $S = \ln(h)$ to express the tax rate as a function of years of schooling S, implying $\frac{\partial t(h)}{\partial h} h = \frac{\partial t(S)}{\partial S}$.[39]

Analogously to the approximation of the linear taxes (and constant tax rates), we can then approximate the simple non-linear tax schedules by regressing the right-hand side of equations (3.41), (3.42), (3.43) for each generation t on a constant and $x = y_t, b_{t+1}, S_{t+1}$, respectively.[40] Given the recursive nature of equation (3.43), we use the approximated taxes in $t+1$ when approximating taxes in t, as before.

The estimated regression coefficients $\widehat{\alpha}_{x,t}$ and $\widehat{\beta}_{x,t}$, for each generation t, then allow us to identify the parameters of interest $\widetilde{\alpha}_{x,t}$ and $\widetilde{\beta}_{x,t}$, $x = y, b, S$ with the following system of equations:

$$\widehat{\alpha}_{y,t} = \widetilde{\alpha}_{y,t}, \tag{3.47}$$

$$\widehat{\alpha}_{b,t} = \widetilde{\alpha}_{b,t}, \tag{3.48}$$

$$\widehat{\alpha}_{S,t} = \widetilde{\alpha}_{S,t} + \widetilde{\beta}_{S,t} \frac{1}{\varsigma_1}, \tag{3.49}$$

$$\widehat{\beta}_{y,t} = 2\widetilde{\beta}_{y,t}, \tag{3.50}$$

$$\widehat{\beta}_{b,t} = 2\widetilde{\beta}_{b,t}, \tag{3.51}$$

[39]See Section 3.3 for an explanation why $S = \ln(h)$ in our model.

[40]The approximation of the linear tax with the cross-sectional average obtains if we only regress the right-hand sides on a constant.

$$\widehat{\beta}_{S,t} = \widetilde{\beta}_{S,t}. \tag{3.52}$$

Bibliography

Abrams, D., Akcigit, U., and Grennan, J. (2018). Patent value and citations: Creative destruction or strategic disruption? NBER Working Paper 19647.

Abrams, D. S., Akcigit, U., and Grennan, J. (2013). Patent Value and Citations: Creative Destruction or Strategic Disruption? Working Paper 19647, National Bureau of Economic Research.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4(B).

Acemoglu, D. and Finkelstein, A. (2008). Input and technology choices in regulated industries: Evidence from the health care sector. *Journal of Political Economy*, 116(5):837–880.

Acemoglu, D. and Linn, J. (2004). Market size in innovation: Theory and evidence from the pharmaceutical industry. *Quarterly Journal of Economics*, 119(3):1049–1090.

Acemoglu, D. and Restrepo, P. (2017a). Robots and jobs: Evidence from us labor markets.

Acemoglu, D. and Restrepo, P. (2017b). Robots and Jobs: Evidence from US Labor Markets. Working Paper 23285, National Bureau of Economic Research.

Acemoglu, D. and Restrepo, P. (2018a). Demographics and automation. NBER working paper 24421.

Acemoglu, D. and Restrepo, P. (2018b). The race between machine and man: Implications of technology for growth, factor shares and employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. U. (2010). When Does Labor Scarcity Encourage Innovation? *Journal of Political Economy*, 118(6):1037–1078.

Adão, R., Kolesár, M., and Morales, E. (2018). Shift-share designs: Theory and inference. *Working Paper*.

Adermon, A., Lindahl, M., and Waldenström, D. (2018). Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. *Economic Journal*, 128:F482–F513.

Aghion, P., Angeletos, G.-M., Banerjee, A., and Manova, K. (2010). Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics*, 57(3):246–265.

Aghion, P., B'enabou, R., Martin, R., and Roulet, A. (2019a). Environmental preferences and technological choices: Is market competition clean or dirty?

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P., and Li, H. (2019b). A theory of falling growth and rising rents.

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., and Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 214(1):1–51.

Aghion, P., Jones, B., and Jones, C. (2017). Artificial intelligence and economic growth. NBER wp 23928.

Akicigit, U. and Ates, S. (2019). Ten facts on declining business dynamism and lessons from endogenous growth theory.

Albanesi, S. and Sleet, C. (2006). Dynamic Optimal Taxation with Private Information. *Review of Economic Studies*, 73(1):1–30.

Alesina, A., Battisti, M., and Zeira, J. (2018). Technology and labor regulations: theory and evidence. *Journal of Economic Growth*, 23(1):41–78.

Arneson, R. (2018). Four Conceptions of Equal Opportunity. *Economic Journal*, 128:F152–F173.

Aschhoff, B., Crass, D., Cremers, K., Grimpe, C., Rammer, C., Brandes, F., Diaz-Lopez, F., Klein Woolthuis, R., Mayer, M., and Montalvo, C. (2010). European competitiveness in key enabling technologies. Technical report, Centre for European Economic Research (ZEW) and TNO.

Atkeson, A. and Lucas, R. E. (1992). On Efficient Distribution With Private Information. *Review of Economic Studies*, 59(3):427–453.

Autor, D. and Dorn, D. (2013a). The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, 103(5):1553–1597.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Reenen, J. V. (2019). The Fall of the Labor Share and the Rise of Superstar Firms. Technical report.

Autor, D., Katz, L., and Krueger, A. (1998). Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics*, 113(4):1169–1213.

Autor, D., Levy, F., and Murnane, R. (2003a). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Autor, D. and Salomons, A. (2018). Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share. Working Paper 24871, National Bureau of Economic Research.

Autor, D. H. and Dorn, D. (2013b). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–1597.

Autor, D. H., Levy, F., and Murnane, R. J. (2003b). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Barkai, Simcha (2016). Declining labor and capital shares. Working Paper, Stigler Center Working Papers.

Barro, R. J. and Lee, J. W. (2013). A New Data Set of Educational Attainment in the World, 1950–2010. *Journal of Development Economics*, 104(C):184–198.

Bartel, A., Ichniowski, C., and Shaw, K. (2007). How Does Information Technology Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills. *Quarterly Journal of Economics*, 122(4):1721–1758.

Bartik, T. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, 75(4):352–365.

Bena, J. and Simintzi, E. (2019). Machines could not compete with chinese labor: Evidence from U.S. firms' innovation. mimeo.

Bessen, J. (2016). How computer automation affects occpations: Technology, jobs and skills.

Bingley, P., Cappellari, L., and Tatsiramos, K. (2018). On the Origins of Socio-Economic Inequalities: Evidence from a "Children of Twins" Design. Manuscript.

Blanas, S., Gancia, G., and Lee, S. Y. (2018). Who is afraid of machines?

Blundell, R., Griffith, R., and Van Reenen, J. (1999). Market share, market value and innovation in a panel of british manufacturing firms. *The Review of Economic Studies*, 66(3):529–554.

Boehm, C. E., Flaaen, A., and Pandalai-Nayar, N. (2019). Multinationals, Offshoring and the Decline of U.S. Manufacturing. Working Paper 25824, National Bureau of Economic Research.

Borusyak, K., Hull, P., and Jaravel, X. (2018). Quasi-experimental shift-share research designs.

Botelho, V. (2018). The Structural Shift in the Cyclicality of the Labor Income Share for the United States. Working Paper, Duke University.

Bovenberg, A. L. and Jacobs, B. (2005). Redistribution and Education Subsidies are Siamese Twins. *Journal of Public Economics*, 89(11-12):2005–2035.

Brown, J. R., Coile, C. C., and Weisbenner, S. J. (2010). The Effect of Inheritance Receipt on Retirement. *Review of Economics and Stastistics*, 92:425–434.

Burda, M. and Seele, S. (2016). No role for the hartz reforms? demand and supply factors in the german labor market, 1993-2014. SFB 649 Discussion Paper 2016-010.

Calel, R. and Dechezleprêtre, A. (2016). Environmental policy and directed technical change: Evidence from the european carbon market. *The Review of Economics and Statistics*, 98(1).

Card, D. (1999). *The Causal Effect of Education on Earnings*, volume 3A of *Handbook of Labor Economics*, chapter 30, pages 1801–1863. Elsevier.

Carnegie-Rochester-Conference (2016). Inequality, Institutions and Redistribution. *Journal of Monetary Economics*, 77(C):1–148.

Cesarini, D., Lindqvist, E., Notowidigdo, M., and Östling, R. (2017). The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries. *American Economic Review*, 107(12):3917–3946.

Chetty, R. (2012). Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, 80(3):969–1018.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2013). Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities. *NBER Macroeconomics Annual*, 27(1):1–56.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics*, 129(4):1553–1623.

Clemens, M., Lewis, E., and Postel, H. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. *American Economic Review*, 108(6):1468–87.

Coelli, F., Moxnes, A., and Ulltveit-Moe, K. (2016). Better, faster, stronger: Global innovation and trade liberalization. NBER wp 22647.

Corbae, D., Nardi, M. D., and Lochner, L. (2017). RED Special Issue on Human Capital and Inequality, an Introduction. *Review of Economic Dynamics*, 25:1–3.

Cortes, G. M., Jaimovich, N., and Siu, H. E. (2017). Disappearing Routine Jobs: Who, How, and Why? page 65.

Cunha, F., Heckman, J. J., and Lochner, L. (2006). *Interpreting the Evidence on Life Cycle Skill Formation*, volume 1 of *Handbook of the Economics of Education*, chapter 12, pages 697–812. Elsevier.

De Nardi, M. and Yang, F. (2016). Wealth Inequality, Family Background, and Estate Taxation. *Journal of Monetary Economics*, 77(C):130–145.

de Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., and van Pottelsberghe de la Potterie, B. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, 42(3):720–737.

Dechezleprêtre, A., Ménière, Y., and Mohnen, M. (2017). International patent families: from application strategies to statistical indicators. *Scientometrics*, 111(2):793–828.

Dechezleprêtre, A., Hémous, D., Olsen, M., and Zanella, C. (2019). Automating Labor: Evidence from Firm-level Patent Data. Working Paper.

Dechezleprêtre, A., Ménière, Y., and Mohnen, M. (2017). International patent families: from application strategies to statistical indicators. *Scientometrics*, 111(2):793–828.

Dickens, R., Machin, S., and Manning, A. (1999). The effects of minimum wages on employment: Theory and evidence from britain. *Journal of Labor Economics*, 17(1).

Dolado, J., Kramarz, F., Machin, S., Manning, A., Margolis, D., and Teulings, C. (1996). The economic impact of minimum wages in europe. *Economic Policy*, 11(23).

Doms, M., Dunne, T., and Troske, K. (1997). Workers, Wages, and Technology. *Quarterly Journal of Economics*, 112(1):253–290.

Eaton, J. and Kortum, S. (1996). Trade in ideas: patenting and productivity in the oecd. *Journal of International Economics*, 40:251–278.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: evidence from french firms. *Econometrica*, 79(5):1453–1498.

Eden, M. and Gaggl, P. (2018). On the welfare implications of automation. *Review of Economic Dynamics*, 29:15–43.

Elsby, M., Hobjin, B., and Sahin, A. (2013). The Decline of the U.S. Labor Share. *Brookings Papers on Economic Activity*, 44(2):1–63.

Farhi, E., Sleet, C., Werning, I., and Yeltekin, S. (2012). Non-linear Capital Taxation Without Commitment. *The Review of Economic Studies*, 79(4):1469–1493.

Farhi, E. and Werning, I. (2007). Inequality and Social Discounting. *Journal of Political Economy*, 115:365–402.

Farhi, E. and Werning, I. (2010). Progressive Estate Taxation. *The Quarterly Journal of Economics*, 125(2):635–673.

Farhi, E. and Werning, I. (2013). Insurance and Taxation over the Life Cycle. *Review of Economic Studies*, 80(2):596–635.

Fernández-Villaverde, J. and Krueger, D. (2007). Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data. *The Review of Economics and Statistics*, 89(3):552–565.

Findeisen, S. and Sachs, D. (2016). Education and Optimal Dynamic Taxation: the Role of Income-Contingent Student Loans. *Journal of Public Economics*, 138(C):1–21.

Findeisen, S. and Sachs, D. (2018). Education Policies and Taxation without Commitment. *The Scandinavian Journal of Economics*, 120(4):1075–1099.

Gaggl, P. and Eden, M. (2018). On the welfare implications of automation. *Review of Economic Dynamics*, 29:15–43.

Gaggl, P. and Wright, G. (2017). A short-run view of what computers do: Evidence from a uk tax incentive. *American Economic Journal: Applied Economics*, 9(3):262–94.

Gentzkow, M. and Shapiro, J. (2010). What drives media slant? evidence from u.s. daily newspapers. *Econometrica*, 78(1):35–71.

Goldsmit-Pinkham, P., Sorkin, I., and Switf, H. (2019). Bartik instruments: What, when, why, and how.

Graetz, G. and Michaels, G. (2017). Is Modern Technology Responsible for Jobless Recoveries? *American Economic Review*, 107(5):168–173.

Graetz, G. and Michaels, G. (2018a). Robots at work. *Review of Economics and Statistics*, 100(5):753–768.

Graetz, G. and Michaels, G. (2018b). Robots at Work. *The Review of Economics and Statistics*, 100(5):753–768.

Güell, M., Pellizzari, M., Pica, G., and Rodríguez Mora, J. V. (2018). Correlating Social Mobility and Economic Outcomes. *Economic Journal*, 128:F353–F403.

Habakkuk, J. (1962). *American and British Technology in the Nineteenth Century*. Cambridge University Press.

Hanlon, W. (2015). Necessity is the mother of invention: Input supplies and directed technical change. *Econometrica*, 83(1):67–100.

Harhoff, D., Scherer, F., and Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy*, 32:1343–1363.

Hausman, J., Hall, B., and Griliches, Z. (1984). Econometric models for count data with an application to the patents-r&d relationship. *Econometrica*, 52(4):909–938.

Heathcote, J., Storesletten, K., and Violante, G. L. (2008). The Macroeconomic Implications of Rising Wage Inequality in the United States. NBER Working Papers 14052, National Bureau of Economic Research, Inc.

Heathcote, J., Storesletten, K., and Violante, G. L. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4):681–722.

Heathcote, J., Storesletten, K., and Violante, G. L. (2014). Consumption and Labor Supply with Partial Insurance: An Analytical Framework. *American Economic Review*, 104(7):2075–2126.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Hémous, D. and Olsen, M. (2018). The rise of the machines: Automation, horizontal innovation and income inequality. CEPR dp 10244.

Henderson, R. and Cockburn, I. (1996). Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery. *The RAND Journal of Economics*, 27(1):32–59.

Henderson, R., Jaffe, A., and Trajtenberg, M. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108(3):578–98.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2008). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *The B.E. Journal of Economic Analysis & Policy*, 7(2):1–48.

Hingley, P. and Park, W. G. (2017). Do business cycles affect patenting? Evidence from European Patent Office filings. *Technological Forecasting and Social Change*, 116:76–86.

Hintermaier, T. and Koeniger, W. (2010). The Method of Endogenous Gridpoints with Occasionally Binding Constraints among Endogenous Variables. *Journal of Economic Dynamics and Control*, 34(10):2074–2088.

Hintermaier, T. and Koeniger, W. (2011). On the Evolution of the US Consumer Wealth Distribution. *Review of Economic Dynamics*, 14(2):317–338.

Holtz-Eakin, D., Joulfaian, D., and Rosen, H. S. (1993). The Carnegie Conjecture: Some Empirical Evidence. *The Quarterly Journal of Economics*, 108(2):413–435.

Hornbeck, R. and Naidu, S. (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review*, 104(3):963–990.

Huggett, M., Ventura, G., and Yaron, A. (2011). Sources Of Lifetime Inequality. *American Economic Review*, 101(7):2923–2954.

Hémous, D. and Olsen, M. (2018). The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality. SSRN Scholarly Paper ID 2328774, Social Science Research Network, Rochester, NY.

Imbens, G. W., Rubin, D. B., and Sacerdote, B. I. (2001). Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players. *American Economic Review*, 91(4):778–794.

Jacobi, L. and Kluve, J. (2007). Before and after the hartz reforms: The performance of active labour market policy in germany. *Journal for Labour Market Research*, 40(1):45–64.

Kapička, M. (2013). Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach. *Review of Economic Studies*, 80(3):1027–1054.

Kapička, M. (2015). Optimal Mirrleesean Taxation in a Ben-Porath Economy. *American Economic Journal: Macroeconomics*, 7(2):219–248.

Kapička, M. (2017). Quantifying the Welfare Gains from History Dependent Income Taxation. Manuscript.

Kapička, M. and Neira, J. (2019). Optimal Taxation with Risky Human Capital. *American Economic Journal: Macroeconomics*, 11(4):271–309.

Karabarbounis, L. and Neiman, B. (2014). The Global Decline of the Labor Share. *The Quarterly Journal of Economics*, 129(1):61–103.

Keane, M. P. and Wolpin, K. I. (1997). The Career Decisions of Young Men. *Journal of Political Economy*, 105(3):473–522.

Kehrig, M. and Vincent, N. (2018). The Micro-Level Anatomy of the Labor Share Decline. Working Paper 25275, National Bureau of Economic Research.

Keynes, J. M. (1930). Economic Possibilities for our Grandchildren. In *Essays in Persuasion*, pages 358–373. Harcourt Brace, 1932, New York.

Kindermann, F., Mayr, L., and Sachs, D. (2018). Inheritance Taxation and Wealth Effects on the Labor Supply of Heirs. Manuscript.

Kocherlakota, N. R. (2010). *The New Dynamic Public Finance*. Princeton University Press.

Koeniger, W. and Prat, J. (2018). Human Capital and Optimal Redistribution. *Review of Economic Dynamics*, 27:1–26.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132(2):665–712.

Koh, D., Santaeulalia-Llopis, R., and Zheng, Y. (2018). Labor Share Decline and Intellectual Property Products Capital. Technical Report 873, Queen Mary University of London, School of Economics and Finance.

Krause, M. and Uhlig, H. (2012). Transitions in the german labor market: Structure and crisis. *Journal of Monetary Economics*, 59(1):64–79.

Krebs, T. and Scheffel, M. (2013). Macroeconomic evaluation of labor market reform in germany. *IMF Economic Review*, 61(4).

Krebs, T. and Scheffel, M. (2017). Labor market institutions and the cost of recessions. IMF working paper WP/17/97.

Krueger, D. and Ludwig, A. (2016). On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium. *Journal of Monetary Economics*, 77(C):72–98.

Krusell, P., Ohanian, L., Ríos-Rull, J.-V., and Violante, G. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, 68(5):1029–1053.

Lee, D., Van der Klaauw, W., Haughwout, A. F., Brown, M., and Scally, J. (2014). Measuring Student Debt and its Performance. Staff Reports 668, Federal Reserve Bank of New York.

Lee, S. Y. T. and Seshadri, A. (2018). Economic Policy and Equality of Opportunity. *Economic Journal*, 128:F114–F151.

Lee, S. Y. T. and Seshadri, A. (2019). On the Intergenerational Transmission of Economic Status. *Journal of Political Economy*, 127(2):855–921.

Lewis, E. (2011). Immigration, skill mix and capital skill complementarity. *Quarterly Journal of Economics*, 126 (2):1029–1069.

Lordan, G. and Neumark, D. (2018). People versus machines: The impact of minimum wages on automatable jobs. *Labour Economics*, 52(40-53).

Lybbert, T. and Zolas, N. (2014). Getting patents and economic data to speak to each other: An algorithmic links with probabilities approach for joint analyses of patenting and economic activity. *Research Policy*, 43:530–542.

Mann, K. and Puttmann, L. (2018). Benign effects of automation: New evidence from patent texts.

Manuelli, R. and Seshadri, A. (2014). Frictionless technology diffusion: The case of tractors. *American Economic Review*, 104(4):1368–1391.

Martinez, J. (2018). Automation, growth and factor shares. mimeo.

Martinez, J. (2019). Automation, Growth and Factor Shares. 2018 Meeting Papers 736, Society for Economic Dynamics.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics*, 114(3):941–975.

Noailly, J. and Smeets, R. (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management*, 72:15–37.

OECD (2011). *Education at a Glance*. OECD.

OECD (2018). *A Broken Social Elevator? How to Promote Social Mobility*. OECD.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic Mechanism Design: A Myersonian Approach. *Econometrica*, 82(2):601–653.

Peterman, W. (2016). The Effect of Endogenous Human Capital Accumulation on Optimal Taxation. *Review of Economic Dynamics*, 21:46–71.

Phelan, C. (2006). Opportunity and Social Mobility. *Review of Economic Studies*, 73:487–504.

Phelan, C. and Rustichini, A. (2018). Pareto Efficiency and Identity. *Theoretical Economics*, 13:979–1007.

Popp, D. (2002). Induced innovation and energy prices. *The American Economic Review*, 92(1):160–180.

Rognlie, M. (2015). Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity? *Brookings Papers on Economic Activity*, 46(1 (Spring)):1–69.

Stantcheva, S. (2015). Optimal Income, Education, and Bequest Taxes in an Intergenerational Model. NBER Working Papers 21177, National Bureau of Economic Research, Inc.

Stantcheva, S. (2017). Optimal Taxation and Human Capital Policies over the Life Cycle. *Journal of Political Economy*, 125(6):1931–1990.

Timmer, M., Dietzenbacher, E., Los, B., Stehrer, R., and de Vries, G. J. (2015a). An illustrated user guide to the world input-output database: the case of global automotive production. *Review of International Economics*, 23(3).

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and Vries, G. J. d. (2015b). An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production. *Review of International Economics*, 23(3):575–605.

van Pottelsberghe de la Potterie, B. and van Zeebroeck, N. (2008). A brief history of space and time: The scope-year index as a patent value indicator based on families and renewals. *Scientometrics*, 75(2):319–338.

WIPO (2019). WIPO Technology Trends 2019: Artificial Intelligence. Technical report, World Intellectual Property Organization, Generva.

Wolff, E. N. and Gittleman, M. (2014). Inheritances and the Distribution of Wealth or Whatever Happened to the Great Inheritance Boom? *Journal of Economic Inequality*, 12:439–468.

Zator, M. (2019). Digitization and Automation: Firm Investment and Labor Outcomes. SSRN Scholarly Paper ID 3444966, Social Science Research Network, Rochester, NY.

Zeira, J. (1998). Workers, Machines, and Economic Growth. *Quarterly Journal of Economics*, 113(4):1091–1117.

CARLO ZANELLA

March 10, 1990

EDUCATION

University of Zurich *since August 2015*

PhD program in Economics

Stockholm School of Economics *July 2014 – June 2015*

Master of Science in Economics

University of St. Gallen *September 2013 – August 2015*

Master of Arts in Quantitative Economics and Finance

University of St. Gallen *September 2010 – August 2013*

Bachelor of Arts in Economics

WORK EXPERIENCE

Research and Teaching Assistant *August 2016 – June 2019*

Department of Economics

Unversity of Zurich

Research Assistant *July 2013 – June 2016*

Swiss Institute for Empirical Economic Research

University of St. Gallen

Intern as Technical Project Manager *February 2010 – June 2012*

CS2

Sissach

February 2020